

University of Groningen

Discovery of Inhibitors by Combinatorial-Chemistry Approaches

van der Vlag, Ramon

DOI:
[10.33612/diss.146091529](https://doi.org/10.33612/diss.146091529)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2020

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):
van der Vlag, R. (2020). *Discovery of Inhibitors by Combinatorial-Chemistry Approaches*. [Thesis fully internal (DIV), University of Groningen]. University of Groningen. <https://doi.org/10.33612/diss.146091529>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Chapter 1

Analytical Methods in Protein-Templated Dynamic Combinatorial Chemistry

Protein-templated dynamic combinatorial chemistry (DCC) has been applied numerous times in the discovery of novel inhibitors of protein targets. A key-step in DCC is the analysis of the libraries to detect an amplification of one or more selected binder(s). To do so, most modern analytical techniques have been applied. We discuss here examples of the application of DCC and more specifically the analysis performed.

This chapter is adapted from the original publication:

R. van der Vlag and A.K.H. Hirsch, Analytical Methods in Protein-Templated Dynamic Combinatorial Chemistry, in: J.L. Atwood (Ed.), *Comprehensive Supramolecular Chemistry II*, Elsevier, Oxford, **2017**, 5, pp. 487–509. ISBN: 9780128031988

1.1 Introduction

In dynamic combinatorial chemistry (DCC), molecular building blocks are reversibly formed via covalent or noncovalent bonds. With only a few building blocks, a relatively large number of products can be generated reversibly, forming a dynamic combinatorial library (DCL). Given that the reaction between the building blocks is reversible, the members of the DCL continuously exchange to eventually reach a thermodynamic equilibrium. When an external stimulus, such as a protein target, is introduced, the dynamic system re-equilibrates upon selection and non-covalent binding of the library members with the strongest affinity for the target. The specific affinity of each library member is determined by the number, type, and strength of interactions (favorable and unfavorable) it establishes with the protein target. Examples of these intermolecular interactions are van der Waals forces, π - π -stacking, cation- π , dipole-dipole, hydrogen-bonding, and ion-dipole interactions, as well as ion pairing. Ultimately, the best binders of the library are amplified at the expense of library members displaying a lower affinity for the target (Fig. 1). Identification of the best binders circumvents the need for synthesis, purification, and characterization of every individual library member. However, up to date, it remains very difficult to analyze DCLs of large size, preventing DCC from becoming a more widely used tool in drug development.¹⁻³ In this article, analytical methods used in protein-templated DCC will be described. Apart from the general theory, examples of recent developments and reports will be briefly discussed.²

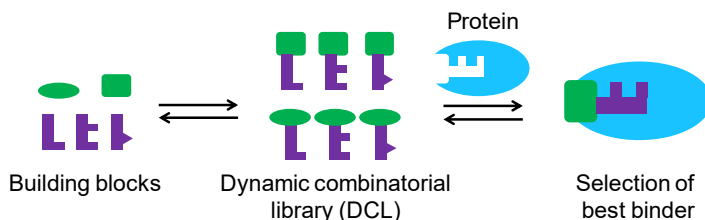
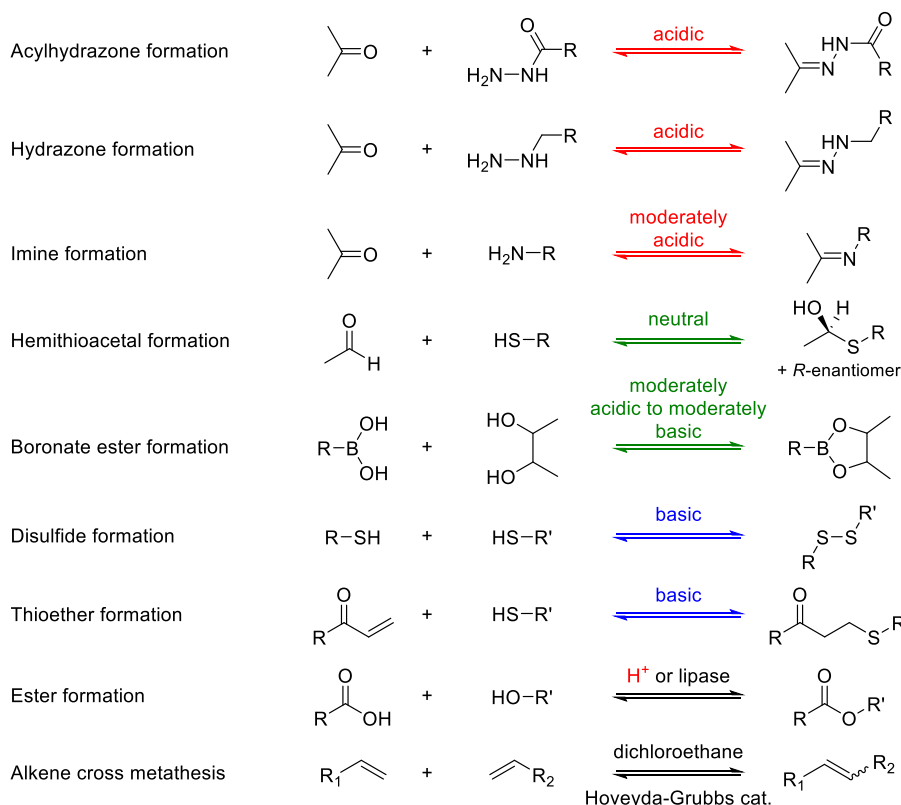


Figure 1. Schematic representation of protein-templated dynamic combinatorial chemistry.

1.2 General Features of DCC Applied to Protein Targets

An overview of the reversible reactions used for the generation of DCLs in protein-templated DCC is given in Scheme 1.^{1,2} Most of these reversible reactions take place in water, which makes them biocompatible. DCL formation can be performed in the presence (adaptive) or in the absence (pre-equilibrated) of the protein target. However, the reversible reaction used for DCL formation should equilibrate fast enough to allow re-equilibration by the target on a timescale during which the target remains stable. The conditions applied should also match the criteria required by the protein, such as a certain pH and temperature and a given concentration of protein, such as a certain pH and temperature and a given concentration of cosolvent(s).



Scheme 1. Reversible reactions used in protein-templated DCC to identify bioactive compounds.²

In order to speed up equilibration, a catalyst such as aniline⁴ or the enzyme lipase⁵ for reversible acylhydrazone or ester formation, respectively, might be added. To ensure formation of an unbiased DCL, the building blocks should be of comparable reactivity and energy. Furthermore, all building blocks and products need to be soluble, to prevent shifts in equilibria due to precipitation. Usually, a cosolvent such as dimethyl sulfoxide (DMSO) is used to maintain products and building blocks in solution. Additionally, it is important that reactions are chemoselective to avoid cross reactivity with functional groups of the building blocks or the protein (*e.g.*, lysine residues at the surface of the protein). Given that the reactions in the DCL are reversible, it should be noted that the analytical method can influence the composition of the library (*e.g.*, trace of acid in liquid chromatography (LC)). To avoid changes in equilibrium composition, libraries should first be “frozen” in the presence of the target protein, whereupon analysis can take place. The method of choice to “freeze” a DCL depends on the type of reaction that is used. Highly reversible imines, for example, can be reduced to the corresponding amines, rendering them irreversible. Acylhydrazones are easily formed and highly reversible under acidic conditions; as a result, the reversible reaction can be slowed down or stopped using an increase in pH. When a reaction is catalyzed, removal or deactivation of the catalyst will convert the DCL into a

static mixture. In each case, carefully designed control experiments need to be carried out, to ensure an unbiased result. In case a reduction step is used to “freeze” the library, it should be studied in how far the binding pose and affinity are affected by this chemical modification. An amine-based inhibitor, for instance, has a lot more flexibility than its imine-based parent compound and an inverted hydrogen-bonding profile. Besides “freezing” the library, the parameters used for DCL generation will affect the equilibrium composition and amplification factors. Examples include the choice of building block and protein concentration, concentration and type of cosolvent (typically 1–10% DMSO), and temperature (typically room temperature). More details can be found in a recently reported review by Hartman *et al.*⁶

Analytical techniques for the analysis of DCLs can be divided into chromatographic and spectroscopic techniques. Often, techniques are combined to obtain the benefits from both, for example, separating a complex mixture with LC and characterizing each member of the library using mass spectrometry (MS).

1.3 General Sample Preparation

Prior to analysis, the sample has to be prepared. After “freezing” the equilibrium, the enzyme can be removed from the sample. Denaturation of the protein is achieved by heating or adding an organic solvent such as acetonitrile or methanol. The denatured protein is subsequently removed by filtration or centrifugation. In the case of MS, it is not always necessary to remove the protein, and a direct analysis of protein–ligand complexes is possible. An overview of the steps in protein-templated DCC is given in Fig. 2.

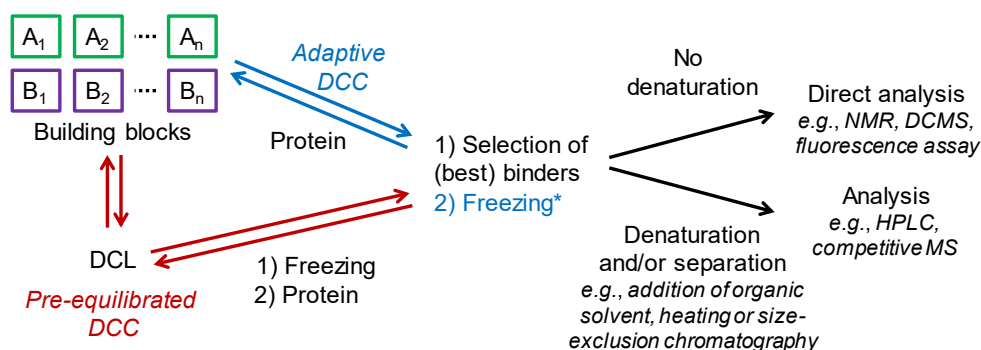


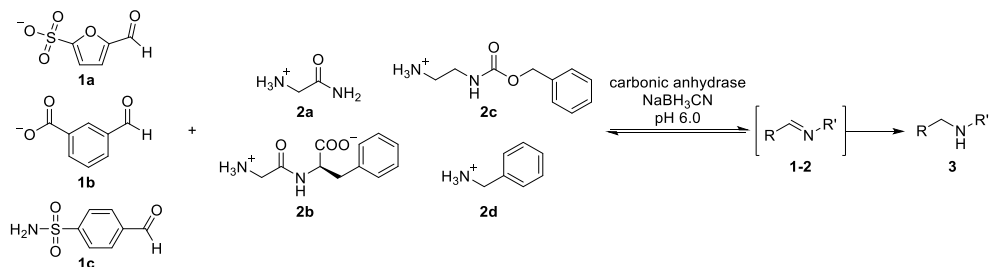
Figure 2. General flow scheme of protein-templated dynamic combinatorial chemistry for the identification of new inhibitors. *Freezing step at this position only applicable for adaptive DCC.

1.4 Analytical Techniques

1.4.1 Chromatographic Techniques

1.4.1.1 Liquid chromatography

Nowadays, access to chromatographic techniques such as high-pressure LC (HPLC) is standard in the modern chemistry laboratory. Owing to the large variety of columns available and a built-in diode array detector, mixtures of compounds can be separated and characterized. Subsequently, the retention time and UV profile of the HPLC chromatogram are analyzed and compared with experiments performed under different conditions. In the case of a DCL, comparison of the library in presence and absence of the target protein can lead to the identification of amplified and/or depleted species. In 1997, DCC was first applied to a protein target by the group of Lehn using imine formation/exchange.⁷ A DCL of three aromatic aldehydes **1a–c** and four different amines **2a–d** in ten-fold excess was synthesized with and without carbonic anhydrase (CA) II as a model enzyme (0.4 mM) (Scheme 2). In order to reduce the imines **1–2** to amines **3**, NaBH₃CN (1.2 mM) was added to the reaction mixture, and the mixture was incubated at 25 °C for 14 days. The reaction without enzyme was completed within 24 h. Before analysis of the DCL, thermal denaturation of the enzyme (2 min. at 80 °C) ensured release from the active site of possibly tightly bound ligands. Using a cellulose membrane, the reaction products were separated from CA by microcentrifuge filtration. Analysis of this first example of protein-templated DCC was performed by comparison of HPLC traces of the DCL in the presence and absence of CA. It was shown that the presence of the enzyme favored the formation of those condensation products that were expected to display the strongest affinity for the enzyme's active site. The strongest binders were identified by comparison with retention times of separately synthesized reference compounds.

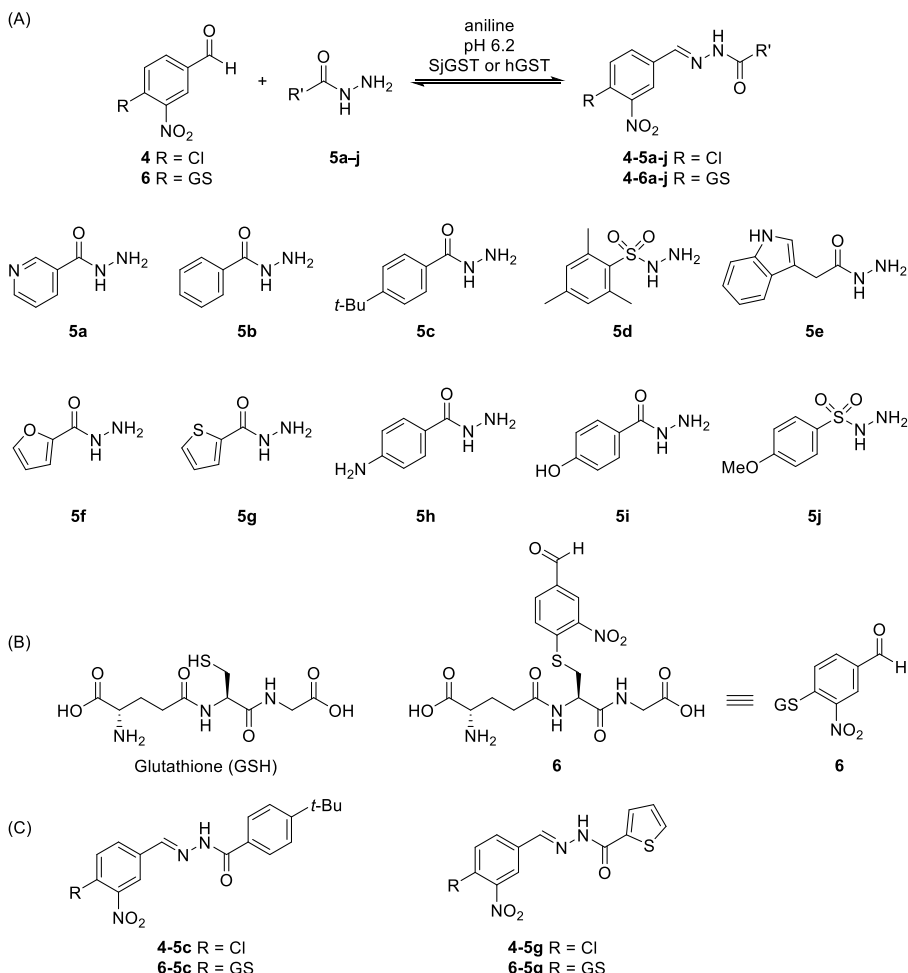


Scheme 2. Dynamic combinatorial library (DCL) library formation with aldehydes **1a–c** and amines **2a–d** in the presence of carbonic anhydrase (CA) results in imines **1–2**, which are subsequently irreversibly reduced to amines **3**.⁷

A more recent example of analysis by LC was reported by Greaney and coworkers,⁸ who had demonstrated in 2010 the use of aniline as a nucleophilic catalyst for the reversible formation of acylhydrazones.⁴ Prior to this report, acylhydrazone chemistry required acidic pH, which is not compatible with most protein targets. Besides a pre-equilibrated system, in which the DCL is formed first, followed by a “freezing” step and addition of the protein, acylhydrazone chemistry was considered to be of limited use for protein-templated DCC.

With the introduction of aniline as a nucleophilic catalyst, acylhydrazones can be formed reversibly at pH 6.2 and reach a steady-state composition in 6 h instead of 7 days without aniline. By controlling the pH, the equilibrium of the DCLs can be easily switched on and off. In the proof-of-principle study, two isozymes of glutathione S-transferase (GST) were used, which play an important role in cell detoxification⁹ and are emerging targets for the treatment of drug resistance in chemotherapy (human (h) GST P1-1)^{10,11} and tropical diseases (*Schistosoma japonicum* (Sj) GST).¹² DCLs were generated with aldehyde **4**, related to the known GST substrate chloro-2,4-dinitrobenzene (CDNB), and a 2.5-fold excess of each of the ten hydrazides **5a–j** in the presence of 10 mM aniline at pH 6.2 (Scheme 3). The hydrazides were added in large excess with respect to the aldehyde to ensure pseudofirst-order behavior and to achieve faster equilibration rates. The DCLs containing **4**, **5a–j**, and one of the isozymes were allowed to stand at room temperature, with occasional shaking, for 12 h. Then, the pH was adjusted to 8.0 to “freeze” the DCL, and the protein was removed by ultrafiltration using a molecular weight cutoff filter of 10,000 Da. Subsequently, the library was analyzed by HPLC and compared with the DCL in the absence of protein. Both DCLs showed a very clear amplification of acylhydrazones: t-butylphenyl and thiophenyl acylhydrazones **4-5c** and **4-5g** were amplified in presence of hGST P1-1 and SjGST, respectively. Accurate IC₅₀ values of the resynthesized acylhydrazones could not be determined due to poor solubility. To solve this problem and at the same time increase the potency of the acylhydrazones, glutathione (GSH)-conjugated aldehyde **6** was synthesized and employed in a DCL with hydrazides **5a–j** (Scheme 3).

The highly soluble GSH tripeptide motif was expected to function as an “anchor” at the GSH-binding site, enabling exploration of the adjacent hydrophobic site with a range of hydrazide fragments. Again, the same hydrazide fragment was selected by the enzymes as the best binder: t-butylphenyl and thiophenyl acylhydrazones **6-5c** and **6-5g** for hGST P1-1 and SjGST, respectively. Acylhydrazone **6-5j** was also slightly amplified in the DCL with hGST P1-1. The most significant reductions in equilibrium concentration occurred for **6-5b**, **f**, and **i** (SjGST) and **6-5f**, **g**, and **i** (hGST P1-1). To confirm that the observed amplification was not caused by target-accelerated synthesis, SjGST was added to a preformed DCL. The same equilibrium distribution was obtained, indicating that the amplification is a result of genuine thermodynamic selection. To fully explore the isozyme-specific amplification of the two DCLs, the IC₅₀ values of the separately synthesized acylhydrazone conjugates **6-5a–j** were determined. The most amplified acylhydrazones were found to be the most potent in the biochemical assay. Thiophenyl acylhydrazone **6-5g** displayed the lowest IC₅₀ value against SjGST (22 μM),

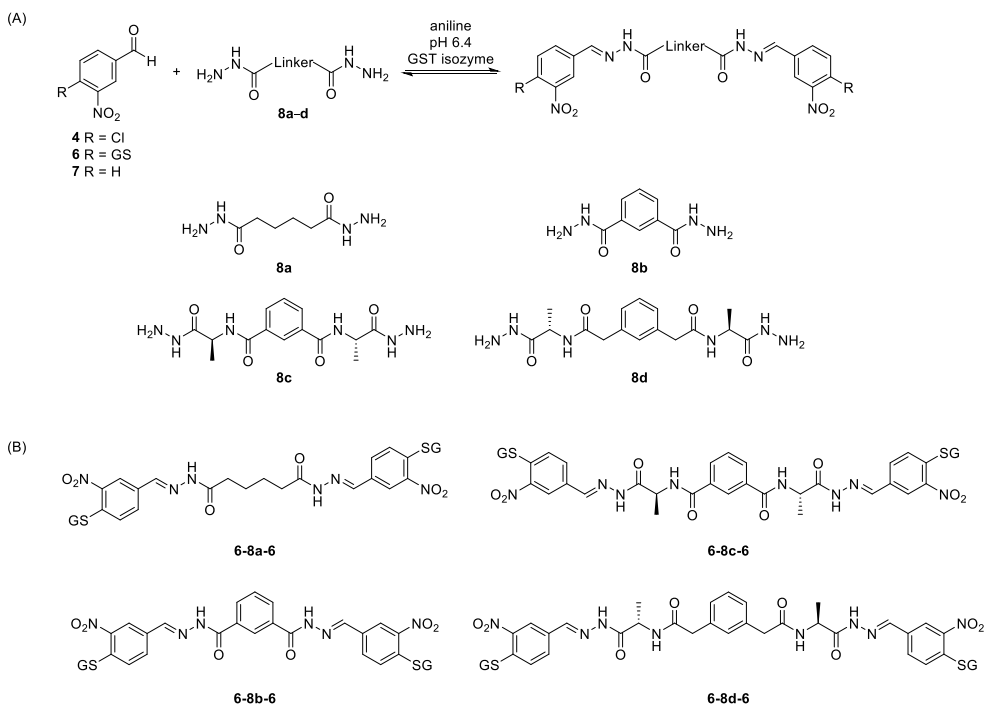


Scheme 3. Generation of acylhydrazone-based dynamic combinatorial libraries of (A) aldehydes **4** or **6** and hydrazides **5a-j**. (B) Structure of glutathione (GSH) and GSH-conjugated aldehyde **6**. (C) Acylhydrazones that were amplified in the presence of GST isoforms *Schistosoma japonicum* (Sj) GST or human (h) GST P1-1.⁴

and t-butylphenyl acylhydrazone **6-5c** had the lowest IC_{50} value among the four conjugates (**6-5c**, **g**, **h**, and **i**) against hGST P1-1 (57 μM).

Based on this initial study, Greaney and coworkers developed stronger, bivalent acylhydrazone inhibitors of the GST isoforms.⁸ This bivalent approach represents a novel concept in protein-templated DCC and provides access to much more (diverse) compounds than in a one-dimensional DCL using the same number of building blocks. The majority of GSTs exist as homodimers, containing a conserved GSH-binding site and a hydrophobic substrate-binding site at the dimer interface. Inhibitors featuring a long “arm” at either end of the central scaffold could be able to span the dimer interface and address both GSH-binding sites. An acylhydrazone-based DCL was generated comprising three nitro-substituted benzaldehydes **4**, **6**, and **7**, which are derivatives of known binder CDNB, and

four bivalent hydrazide linkers **8a-d** in the presence of aniline as a nucleophilic catalyst (Scheme 4A), resulting in 24 homo- and hetero-bis-acylhydrazones (plus additional mono-acylhydrazones), excluding *E/Z* isomers. In order to have the opportunity to discriminate between the H-sites of different GST homodimers, the linkers were designed to have differing lengths and lipophilicity. Four GST isozymes (murine (m) GSTM1-1, hGSTP 1-1, SjGST, and mGSTA4-4) were used as template for the reversible formation of acylhydrazone binders. After incubation at room temperature for 48 h, HPLC analysis of the adaptive DCLs using two HPLC columns in series was performed and compared with DCLs generated in the absence of protein. Each peak was identified by deconvolution in combination with LC-MS analysis. The highest amplifications were obtained for the m and Sj isozymes, with an impressive amplification of over 600% by mGSTM1-1 for hydrazone **6-8c-6** (Scheme 4B).



Scheme 4. (A) Aniline-catalyzed generation of acylhydrazone-based dynamic combinatorial libraries of aldehydes **4**, **6**, **7** and bis-hydrazides **8a-d**. (B) Amplified GSH-conjugated homo-bis-acylhydrazones in the presence of mGSTM1-1.⁸

Table 1. Table of IC₅₀ values of amplified bis-acylhydrazones **6-8-6** against four different GST isozymes.⁸

Acylhydrazone	IC ₅₀ (μM)			
	mGSTM1-1	hGSTP1-1	SjGST	mGSTA4-4
Aldehyde 6	341.7	≥ 500	265.6	> 500
Hydrazide 8c	> 500	> 500	ND	ND
6-8a-6	1.207	126.5	3.471	> 100
6-8b-6	0.337	11.81	0.252	> 100
6-8c-6	0.050	13.45	0.989	> 100
6-8d-6	0.413	0.356	1.800	> 100

ND = Not determined

The DCL containing mGSTM1-1 showed strong amplification of acylhydrazones formed from GSH-conjugated aldehyde **6**. The presence of mGSTA4-4 did not lead to significant changes in the HPLC traces and was later used as negative control to study specific binding-site interactions within the mGSTM1-1 isoform. In contrast, hGSTP1-1 appeared to interfere with the DCL equilibrium for certain compounds. For example, bis-acylhydrazones containing two units of **6** were degraded to mono-acylhydrazones by the enzyme. Therefore, hGSTP1-1 was not used further in the templating studies.

A range of symmetrical bis-acylhydrazones, all containing GSH-conjugated **6**, was selected and separately synthesized in order to assess the corresponding biochemical activities (Scheme 4B and Table 1). For mGSTM1-1, the IC₅₀ data correlate well with the amplifications observed in the DCL: inhibitor **6-8c-6** (IC₅₀ = 50 nM), which showed the greatest amplification, is nearly ten-fold more potent than the other bis-acylhydrazone products. The inhibition data for SjGST correlate less well with the amplification observed, presumably due to competition among the amplified library members for shared building block **8b**, which reduces the extent of amplification. Due to a lack of clear equilibration, the amplification factors observed could not be correlated with the IC₅₀ values for hGSTP1-1. To elucidate the role of linker **8c**, an overlay model was constructed based on the crystal structure of hGST M1A-1A, containing ligand 5-2,4-dinitrobenzene. The model shows that the linker **8c** can span the dimer interface of hGST M1A-1A, without introducing any unfavorable interactions or distorting the dimer structure, demonstrating that DCC with bivalent linkers is particularly suited for protein targets that are challenging to address by structure-based design (SBD).

At the moment, LC techniques such as HPLC are the most widely used techniques. LC is often combined with MS, in which case peaks can immediately be assigned to the corresponding library member. Otherwise, first, the retention times and UV profiles have to be confirmed using separate building blocks and (side) products. New potential inhibitors that do not contain a strong absorbing chromophore can be overlooked when only UV detection is used. Evaporative light scattering detectors (ELSD) might solve this problem. However, the use of ELSD requires different equipment and is therefore less commonly used. One of the major bottlenecks associated with HPLC and LC–MS is the limited library size. Although the analytical method possesses adequate sensitivity to detect subtle

changes in concentration of compounds of interest, a mixture of many compounds can often not be fully resolved. This limits the size of the DCL to tens of compounds rather than thousands of possible ligands. In the future, it is expected that methods using ultra performance LC (UPLC) will increase the potential library sizes, but even then, there will be a size limit. An additional downside of HPLC analysis is the requirement for relatively large amounts of protein.

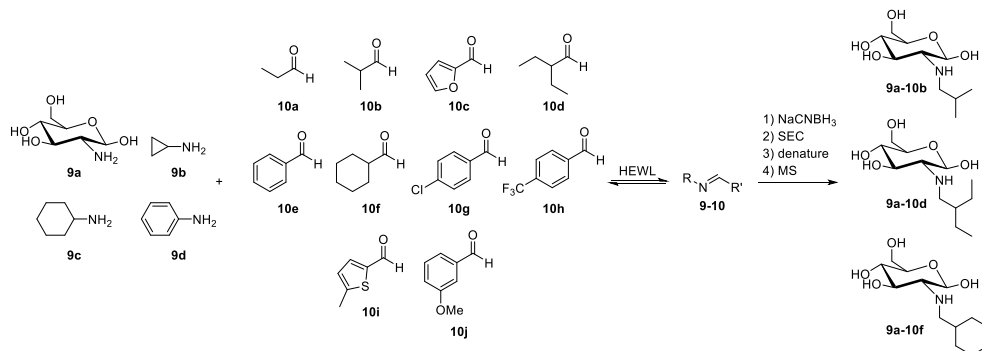
1.4.1.2 Size-exclusion chromatography

In 2013, the group of Guo reported a novel protocol based on size-exclusion chromatography (SEC) and MS for direct isolation and identification of ligand–target adducts from DCLs.¹³ For the proof-of-concept study, hen egg white lysozyme (HEWL) was chosen as a protein target, which is an important contributor to immune defense by attacking and degrading bacterial cell walls through hydrolysis.

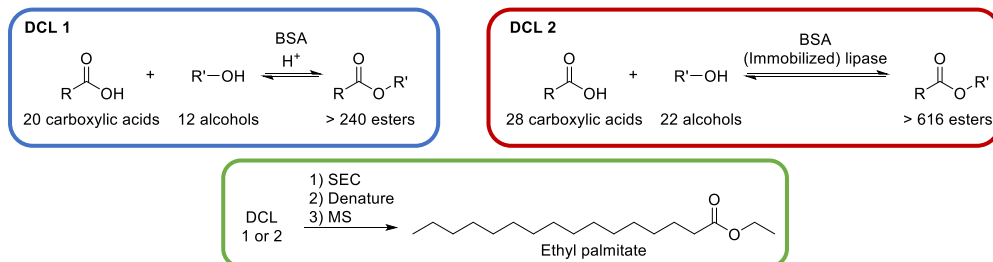
Imine formation was selected as the reversible reaction, and a DCL of four amines **9a–d** and ten aldehydes **10a–j** was designed (Scheme 5). The adaptive DCL was reduced to the corresponding amines using NaCNBH_3 . After incubation with HEWL, the library mixture was passed through an SEC column to retain all the nonbinders in the library. The potential binders were released by denaturation of the eluted protein–ligand complexes using acetonitrile, and the eluent was analyzed by electrospray ionization (ESI)–MS. Three members of the DCL were found to bind HEWL. In control reactions in the absence of protein no compounds were detected by MS, indicating that the SEC column could retain all the unbound library members and that amines **9a–10b**, **9a–10d**, and **9a–10f** are indeed binders of HEWL. In case the library was not reduced to the corresponding amines, no imines were detected after SEC, suggesting that amines rather than imine intermediates are the ligands of HEWL. The amines were synthesized, and their mode of inhibition and relative potencies were determined by evaluating their effect on the lysis rate of *Micrococcus lysodeikticus*.¹⁴ The amines were found to adopt a competitive mode of inhibition, and the binding affinities follow the order **9a–10f** ($K_m = 0.203 \pm 0.028 \text{ mg mL}^{-1}$) > **9a–10b** ($K_m = 0.163 \pm 0.022 \text{ mg mL}^{-1}$) \approx **9a–10d** ($K_m = 0.162 \pm 0.063 \text{ mg mL}^{-1}$).

In 2015, building on their first report, Guo and co-workers developed a SEC–MS protocol for the identification of competitive inhibitors for bovine serum albumin (BSA).⁵ Serum albumins play an important role in the transport and delivery of fatty acids, metals, drugs, and other bioactive small molecules and are the most abundant soluble protein constituents of the circulatory system. Two adaptive DCLs (DCL1 and DCL2, Scheme 6) were prepared using different catalysts consisting of 20 carboxylic acids (408 μM each) and 12 alcohols (500 μM) and 28 carboxylic acids (480 μM each) and 22 alcohols (550 μM). Sulfuric acid and an immobilized lipase *Candida* sp. 99–125 were added as a catalyst to DCL1 and DCL2, respectively. After equilibration for three days, aliquots of the solutions were applied to a suitable SEC column. After denaturing the protein by the addition of MeOH and centrifugation, the supernatant was analyzed by ESI–HRMS. Interestingly, the library was not “frozen” before denaturation. From both DCLs, ethyl palmitate was identified as a new binder of BSA. The binding and affinity of the protein for ethyl palmitate were further

examined via fluorescence spectroscopy. The thermodynamic parameters and association constants K_a were determined using the Van 't Hoff and Stern–Volmer equations, respectively. Based on the thermodynamic results, the binding process is spontaneous ($\Delta G < 0$), and ethyl palmitate is considered to be bound to BSA mainly through hydrophobic interactions.



Scheme 5. Generation of a dynamic combinatorial library of amines **9a-d** and aldehydes **10a-j** in the presence of hen egg white lysozyme (HEWL), followed by imine reduction with NaCNBH₃ and separation of protein–binder complexes by size-exclusion chromatography (SEC).¹³



Scheme 6. Dynamic combinatorial libraries 1 (blue box) and 2 (red box) and the SEC–MS protocol for analysis (green box).

SEC–MS protocols circumvent the need for resolution of all library members by LC, which is one of the major bottlenecks, preventing the use of significantly larger library sizes. A drawback of SEC, however, is the fact that the concentration of the building blocks for DCL generation is limited to the micromolar or even nanomolar range. This can result in long equilibration times, which can cause problems with protein stability. Another disadvantage are the large amounts of protein used (*e.g.*, 30 mg HEWL¹³ or 6–10 mg BSA⁵ per DCL) to ensure a sufficiently high concentration of binding inhibitors for proper detection by MS. With enzymes that are difficult to express or unstable over time, the SEC–MS protocol is therefore not suitable yet. This problem could possibly be addressed with selected-ion monitoring MS, in which case selected m/z values are detected, but this will limit the library size or increase the analytical time significantly.

1.4.2 Nuclear Magnetic Resonance Spectroscopy

Nuclear magnetic resonance (NMR) spectroscopy is a commonly used tool in organic chemistry laboratories. Since it is nondestructive, it can be performed directly in the reaction mixture and is extremely sensitive, making it an invaluable tool in chemistry. Several NMR-based methods have been developed and used in protein-templated DCC, including saturation-transfer difference (STD)-NMR, ^{11}B -NMR, water-ligand observed via gradient spectroscopy (waterLOGSY), and competition-based ^1H -NMR spectroscopy.

1.4.2.1 ^1H -STD-NMR spectroscopy

In STD-NMR spectroscopy, the difference between a “normal” ^1H -NMR spectrum and an ^1H -NMR spectrum, in which the protein is selectively saturated, is studied (Fig. 3).^{15,16} First, a reference spectrum is recorded with the irradiation frequency set at a value far away from any ligand or protein signal. This spectrum is also known as the off-resonance spectrum. In a second experiment, a spectrum is recorded while selectively irradiating the protein, without influencing the ligand(s). Usually, this saturation step is performed between 0.0 and -1.0 ppm, a region characteristic of aliphatic residues of the protein. The selective saturation is now transferred from the aliphatic side chains to the whole protein via spin diffusion through cross relaxation (intramolecular nuclear Overhauser effect). Fast exchange takes place between the free and bound ligands, which allows further transfer of magnetization from the protein receptor to the ligand protons, which are involved in binding to the protein’s surface. The saturation that is transferred to the ligands is detected as a change in ligand signals in the saturated STD-NMR spectrum. Upon subtraction of the on-resonance spectrum from the off-resonance spectrum, only signals are observed from ligands that are in close contact with the protein.

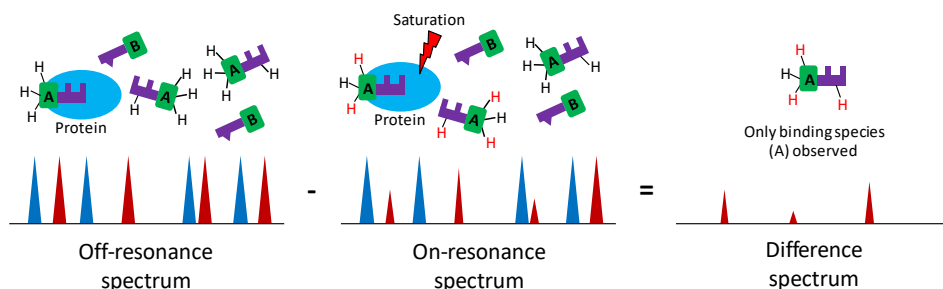
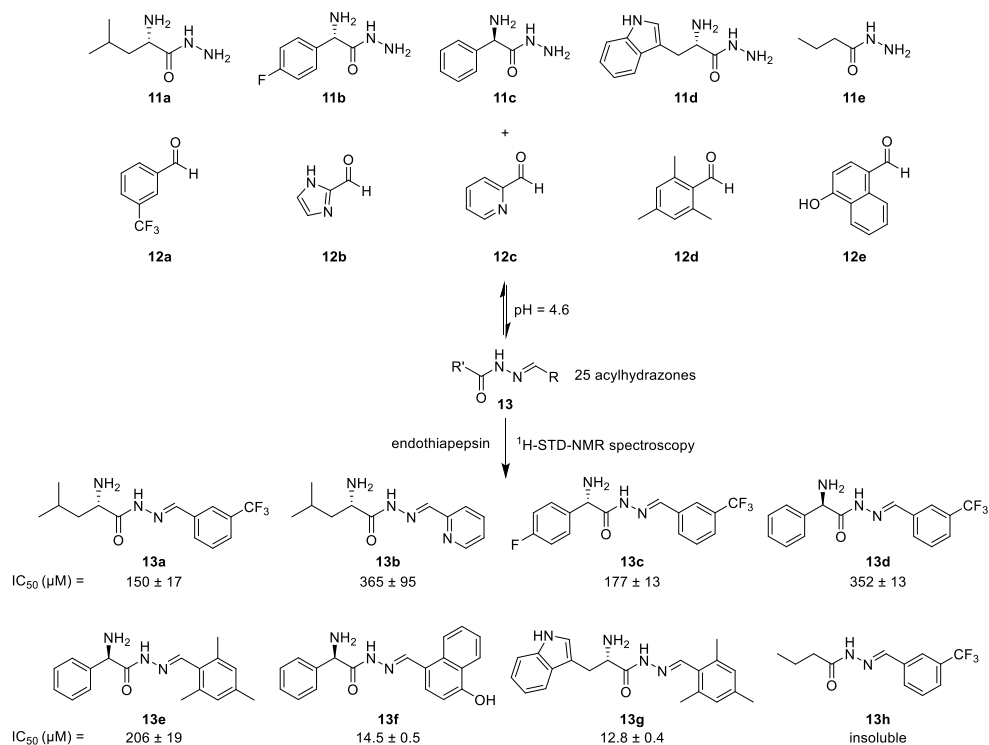


Figure 3. Schematic representation of the ^1H -STD-NMR experiment. Intramolecular magnetization is transferred from the protein to the binders (A) by the exchange of free and bound ligand. As a result, nonbinders (B) are not observed in the last spectrum. Adapted from Viegas, A.; Manso, J.; Nobrega, F. L.; Cabrita, E. J.; *J. Chem. Educ.* **2011**, *88* (7), 990–994.

In 2014, we used *de novo* SBD in combination with acylhydrazone-based DCC and ^1H -STD-NMR spectroscopy to identify new inhibitors of the aspartic protease endothiapepsin.¹⁷ The family of aspartic proteases is involved in a wide range of diseases such as malaria (plasmeprins), Alzheimer's disease (β -secretase), and hypertension (renin).¹⁸ Based on the interactions observed in two crystal structures of endothiapepsin (with and without a water molecule in the active site), a series of acylhydrazones were designed. Retrosynthetic analysis of the acylhydrazones resulted in five hydrazides **11a–e** and five aldehydes **12a–e** (Scheme 7). To facilitate analysis by ^1H -STD-NMR spectroscopy, the full library was divided into five sublibraries, each consisting of one aldehyde and five hydrazides. After the addition of endothiapepsin to the preformed sublibraries, the acylhydrazones were identified by analyzing the imine-type (**12a**, **12b**, **12d**, and **12e**) or α -carbon (**12c**) proton signals in the ^1H -STD-NMR spectra. From the five sublibraries, featuring a total of 25 potential acylhydrazones **13** (50 isomers including *E/Z* isomers), eight binders were identified (Scheme 7, **13a–h**). In order to confirm the binding mode, the competitive, potent inhibitor saquinavir (inhibition constant (K_i) = 48 nM) was added to the **11a–e + 12d** sublibrary. The appearance of saquinavir signals and decrease of acylhydrazone signals in the ^1H -STD-NMR spectrum confirmed specific binding of the acylhydrazones **13e** and **13g**. Subsequently, the biochemical activity of the eight binders was determined using a fluorescence-based assay. Seven of the eight acylhydrazones inhibit endothiapepsin with IC_{50} values between 13 and 365 μM (Scheme 7), confirming the outcome of the ^1H -STD-NMR experiments (one acylhydrazone was insoluble under the biochemical assay conditions). The most potent inhibitors **13g** and **13f** display IC_{50} values of 12.8 and 14.5 μM , respectively. In order to determine the predicted binding mode, endothiapepsin crystals were soaked with inhibitors **13g** and **13f**. Cocystal structure determination confirmed the *in silico* prediction that either direct or water-mediated interactions with the catalytic dyad can be achieved.



Scheme 7. Generation of preformed acylhydrazone-based libraries of hydrazides **11a–e** and aldehydes **12a–e**, followed by addition of the protein target endothiapepsin and analysis by ^1H -STD-NMR spectroscopy. IC_{50} values provided by a fluorescence-based enzyme activity assay.¹⁷

^1H -STD-NMR spectroscopy has the benefits that it is nondestructive and can be directly performed using the sample. Furthermore, since the STD method indicates which hydrogen atoms of the small-molecule inhibitors are in close proximity of the protein, it provides some information on the binding mode. In addition, competition experiments with known binders can elucidate if the ligands bind in the same pocket as the known ligand (assuming that the binding mode of the latter is known). STD-NMR spectroscopy only works if the ligand is used in excess, therefore requiring significantly less protein than LC-MS. It should, however, be noted that this technique does not discriminate between specific and nonspecific binders.¹⁹ Calculation of dissociation constants (K_d) is therefore highly overestimated, and additional experiments such as IC_{50} determination are required. High-affinity binders that undergo slow chemical exchange on the NMR timescale are not observed.¹⁶ ^1H -STD NMR has only been applied twice in the analysis of a DCL, first in 2010 by the group of Ramström in the analysis of β -galactosidase inhibitors based on hemithioacetals²⁰ and in 2014 by our group,¹⁷ as discussed earlier.

1.4.2.2 Other NMR-based methods

Besides ^1H -STD-NMR spectroscopy, more NMR methods have been developed for the analysis of protein-templated DCLs. Claridge and Schofield reported the use of two NMR methods to analyze DCLs with proteins as template for the first time, namely, ^{11}B -NMR spectroscopy and ^1H -waterLOGSY.²¹ Using these NMR techniques, ternary complexes of enzyme, boronic acids, and sugars in the DCLs were monitored to identify binders of α -chymotrypsin. The ^{11}B nucleus is quadrupolar ($I = 3/2$), which can be exploited to determine the hybridization state of the boron atom: an ^{11}B nucleus in an sp^2 -hybridized trigonal planar environment exhibits relatively broad peaks compared with an ^{11}B nucleus in a highly symmetrical sp^3 -hybridized tetrahedral environment, which gives much sharper peaks.²² Although ^{11}B -NMR spectroscopy enables the direct observation of the boronate–enzyme complexes and is free from protein background signals, it is far from optimal for the analysis of DCLs. The ^{11}B -NMR signals are broad, causing boronic acids with similar structures to overlap. Additionally, ^{11}B -NMR spectroscopy suffers from low intrinsic sensitivity, which necessitates the use of high concentrations of boronic acids and protein to compensate for the loss in signal. ^1H -waterLOGSY²³ features better sensitivity than ^{11}B -NMR spectroscopy. First, magnetization of the bulk H_2O in the sample takes place, and after a variety of different transfer mechanisms, small molecules that interact with the target protein are detected. The protein concentration can be very low, but ^1H -waterLOGSY can be limited by the exchange kinetics of the reversibly formed protein–ligand complexes and might still suffer from overlapping signals, especially for large DCLs.

In 2013, the same group introduced a competition-based ^1H -NMR method to screen binders for human 2-oxoglutarate (2OG)-dependent oxygenases.²⁴ Hypoxia inducible factor (HIF) hydroxylases, specifically prolyl hydroxylase domain containing enzyme isoform 2 (PHD2) and factor-inhibiting HIF (FIH), were used as model systems for this study. The natural cosubstrate 2OG was used as a reporter ligand, and in order to block enzyme-catalyzed 2OG turnover and to avoid oxidation of diamagnetic Fe(II) , the native Fe(II) was substituted by paramagnetic and catalytically inactive Zn(II) . When a competitive binder displaces bound 2OG from the binding site into the bulk solution, the unbound 2OG is observed via ^1H -NMR spectroscopy. By monitoring the intensity of the reporter signal as a function of inhibitor concentration, it is also possible to obtain dissociation constants (K_d). As a proof-of-principle study, the boronic acid scaffold and some diol hits from a previous study²⁵ using dynamic combinatorial MS (DCMS; see section 1.4.3.1) were subjected to the competitive NMR analysis. The results obtained are in agreement with the nondenaturing DCMS results. The competition-based ^1H -NMR technique only uses substoichiometric amounts of unlabeled protein and has additional advantages over other ligand-based NMR techniques: site-specific binding information and K_d values for ligands with both high and low affinity can be obtained.

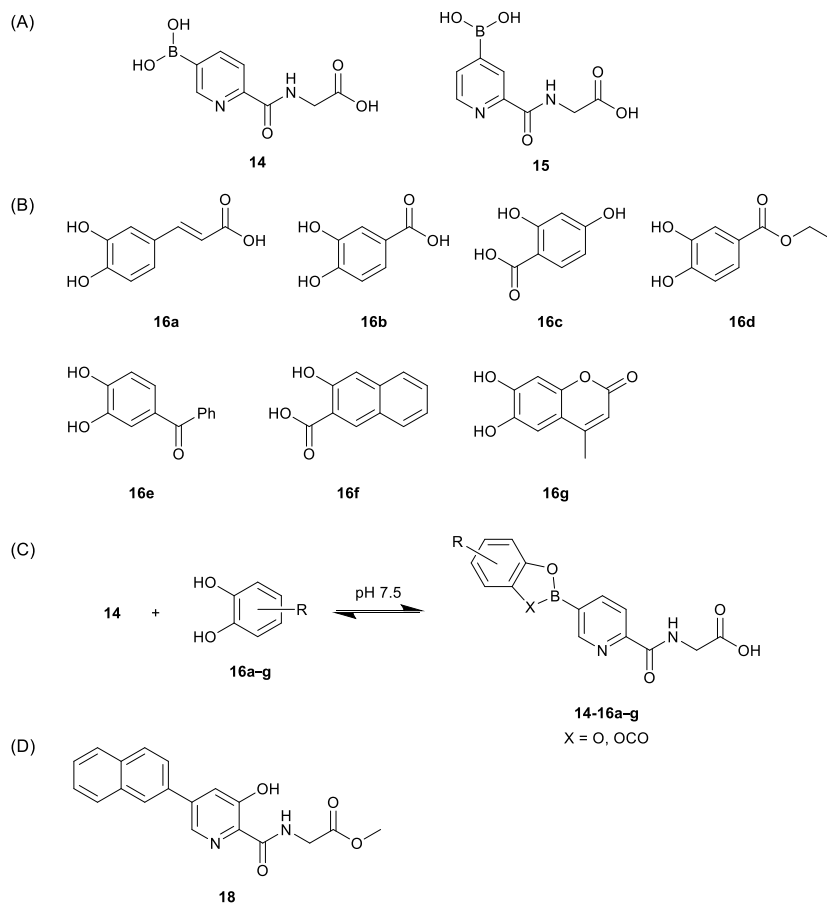
1.4.3 Mass Spectrometry

As already briefly introduced in section “Liquid chromatography,” MS is a very useful tool for DCC, since compounds can immediately be assigned by their mass. Prior to analysis, not

only complex mixtures can be separated using LC but also direct analysis of DCLs can be applied. The two main MS techniques, which are applied to protein-templated DCC, are discussed in the next sections: “Dynamic combinatorial MS” and “Competitive MS”.

1.4.3.1 Dynamic combinatorial MS

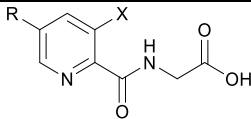



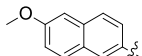
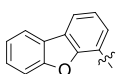

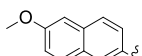
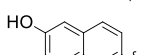
Nondenaturing DCMS was introduced by Poulsen in 2006.²⁶ In the DCL, one of the library members is known to bind strongly to the protein and serves in this way as an “anchor”. Next, the reversible reaction takes place with this “support ligand” to obtain more favorable interactions with the binding site of the enzyme. Subsequently, nondenaturing ESI-MS and MS–MS are used for the direct analysis of protein–ligand complexes. A recent example of DCMS was reported in 2012 by the group of Schofield enabling the identification of new inhibitors based on acids/boronate esters for PHD2 (see section “Other NMR-based methods”).²⁵ PHD2 is an Fe(II) and 2OG oxygenase that regulates human hypoxic response. Involvement in treatment of anemia and ischemia-related diseases makes this human hydroxylase an important drug target.²⁷ First, based on modeling studies, “support ligands” **14** and **15** were designed to participate in both Fe(II) chelation in the active site and boronate ester exchange (Scheme 8A). It was predicted that “support ligand” **14** fits well in a side pocket of the active site and **15** would function as a control, since it should not bind at the active site due to steric clash. Four sublibraries of ten diols each were prepared based on molecular weight, and PHD2–Fe(II) and “support ligands” **14** or **15** were added to each DCL (pH 7.5). Subsequently, the eight DCLs (four containing **14** and four containing **15**) were analyzed with ESI-MS in order to determine which boronate esters bind preferentially to PHD2–Fe(II). Of the 40 possible boronates with “support ligand” **14**, seven boronate esters derived from seven diols **16a–g** were found to bind (Scheme 8B). ESI-MS analysis of mixtures of **14** with the individual diols validated the DCMS results. In the absence of **14**, some of the individual diols (*e.g.*, **16e** and **16g**) were also observed to bind to the enzyme, presumably through chelation with Fe(II). When **15** was employed as a control, no binding of boronate esters was found, and again, binding of some diols to the enzyme was observed. These results indicate that the mass shifts observed with **14** represent the boronate ester–PHD2 complexes, instead of simultaneous binding of **14** and the diols. In order to validate the DCMS results and to ensure that the results obtained were not biased by the MS technique, NMR-based water relaxation experiments (see section “Other NMR-based methods”) were performed, confirming that the DCMS method (gas phase) is in agreement with the NMR studies (solution).



Scheme 8 (A) Support ligands **14** and **15**. (B) Selected diol building blocks **16** from four dynamic combinatorial libraries (DCLs) with **14** and **15**. (C) DCL formation of boronate esters **14-16a-g** from boronic acid **14** and diol building blocks **16a-g**. (D) Structure of methylester derivative **18**.²⁵

Verification of the potential of the boronic acid-/boronate ester-based DCMS method was performed by determining the binding constants of stable analogs **17a-i** of “support ligand” **14** and diols **16a-g** (Scheme 8 and Table 2). All analogs bind significantly stronger to PHD2 than boronic acid **14**, but generally weaker than boronate esters **14-16a-g** identified by DCMS. In order to investigate the binding mode of the boronate esters, stable analog **17h** was crystallized in complex with PHD2, containing Mn(II) as a substitute for Fe(II). The protein cocrystal structure revealed that **17h** binds to the metal in a bidentate fashion. To study its effect in cells, activity studies were performed with methyl ester derivative of **17g** (**18**, Scheme 8). Compound **18** was found to upregulate HIF1a by selectively inhibiting PHDs, but not FIH.

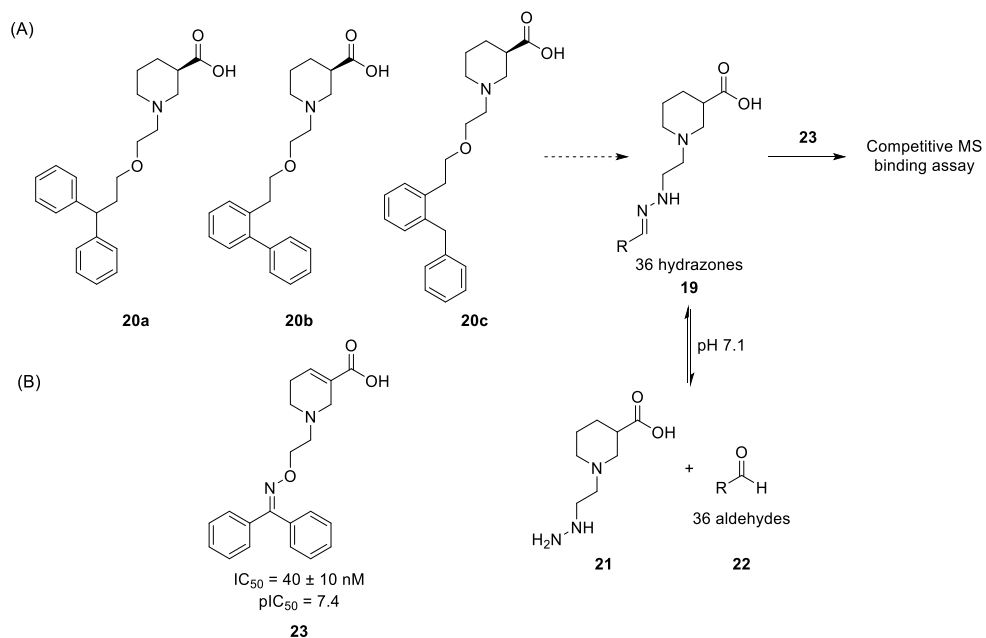
Table 2. Binding constants (K_D) and IC_{50} values for stable analogs **17a–i** of boronic acid inhibitors **14–16a–g**.²⁵

				
Compound	R	X	K_D (μM)	IC_{50} (μM)
14		H	24.8	126
17a	H	H	-	> 1000
17b		H	9.5	> 500
17c		H	7.0	> 500
17d		H	1.6	107
17e		H	8.7	> 100
17f	H	OH	3.5	409
17g		OH	0.5	0.017
17h		OH	0.8	0.013
17i		OH	0.9	0.004

Overall, the authors demonstrated that inhibitors based on boronate esters could be discovered using DCMS in combination with NMR spectroscopy. The DCMS method prevents the need for denaturing the protein, chromatography to separate the reaction mixture, and “freezing” of the DCL. The technique is very fast and enables direct analysis of the protein–ligand complexes. Synthesis of all individual library members to validate spectral assignments is not required, and individual building blocks could be directly used in determining binding constants. Despite the high sensitivity of MS, however, the library of 40 compounds had to be divided into sublibraries of ten possible compounds. Schofield and coworkers also described some fragmentation of boronic acids, which is a limitation of the DCMS approach. Also not all proteins can be used for nondenaturing MS analysis, and during analysis, care should be taken, since not all of the noncovalent interactions might be translated during the transition from solution to gas phase in the same manner.²⁸ Lastly, due to the soft conditions that have to be used, DCMS methods are not that straightforward, and a lot of time and effort can be required to establish a method for analyzing a DCL with a novel protein target.

1.4.3.2 Competitive MS

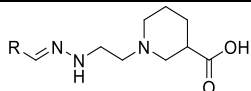
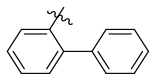
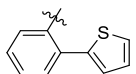
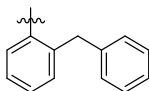
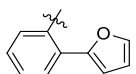
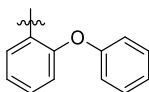
A slightly different approach is the screening of DCLs by competitive MS-based binding assays. In 2012, the group of Wanner demonstrated an efficient way to screen hydrazone inhibitors using pseudostatic libraries and determined the binding affinity by competitive MS binding assays.²⁹ γ -Aminobutyric acid (GABA) transporter 1 (GAT1) was chosen as a target, which is the most important subtype of the GABA transporters. Decreased GABAergic neurotransmission is assumed to be a major cause of neurological disorders like epilepsy, Parkinson's disease, and sleeping disorders.^{30,31} Inspired by the work of Andersen et al.,³² hydrazone derivatives **19** were designed that resemble mGAT1 inhibitors **20a–c** (Scheme 9). Nipecotic acid derivative **21** was used as the hydrazine building block, and 36 aldehydes **22** were divided over nine equal libraries (four possible hydrazones per DCL). In order to get a rough estimate of the equilibration time, five different aromatic aldehydes (40 μ M to resemble four aldehydes at 10 μ M) were reacted separately with hydrazine **21** (100 μ M) in the absence of target, and the reaction was monitored by UV. In each case, almost complete equilibrium was reached within 4 h, and therefore, a 4 h incubation period was chosen for the DCC experiments in the presence of the target (adaptive DCC). To ensure nearly full conversion of the aldehydes toward the products, hydrazine **21** (100 μ M) was used in a 2.5-fold excess with respect to the total aldehyde concentration (four aldehydes, 10 μ M each). In case equal amounts of the structurally diverse aldehydes are added to the hydrazine, a library should be obtained with all constituents being present in similar amounts. Although the building blocks are continuously exchanging, the overall composition of the library is almost constant. Such a library is also known as a pseudostatic DCL. After addition of only 10–20 μ g of protein per sample and incubation at 37 °C for 4 h, the DCLs were analyzed by competitive MS. For hit detection, native marker **23** (K_d 26.5 \pm 4.6 nM at pH 7.1) was added to each sample (final concentration **23** 20 nM) and allowed to equilibrate for 40 min. After termination of the reaction by vacuum filtration and denaturation of the isolated protein–ligand complexes, the amount of MS marker was quantified by LC–ESI–MS/MS. Control libraries containing no hydrazine **21**, and hydrazine **21** without aldehydes were shown to have minor effects on **23** binding, decreasing it to a maximum of 70% and 92 \pm 3%, respectively.



Scheme 9. (A) GAT1 inhibitors **20a–c** as inspiration for DCLs of hydrazones **19** using hydrazine **21** and 36 aldehydes **22**. (B) Known GAT inhibitor **23** as MS-binding marker.²⁹

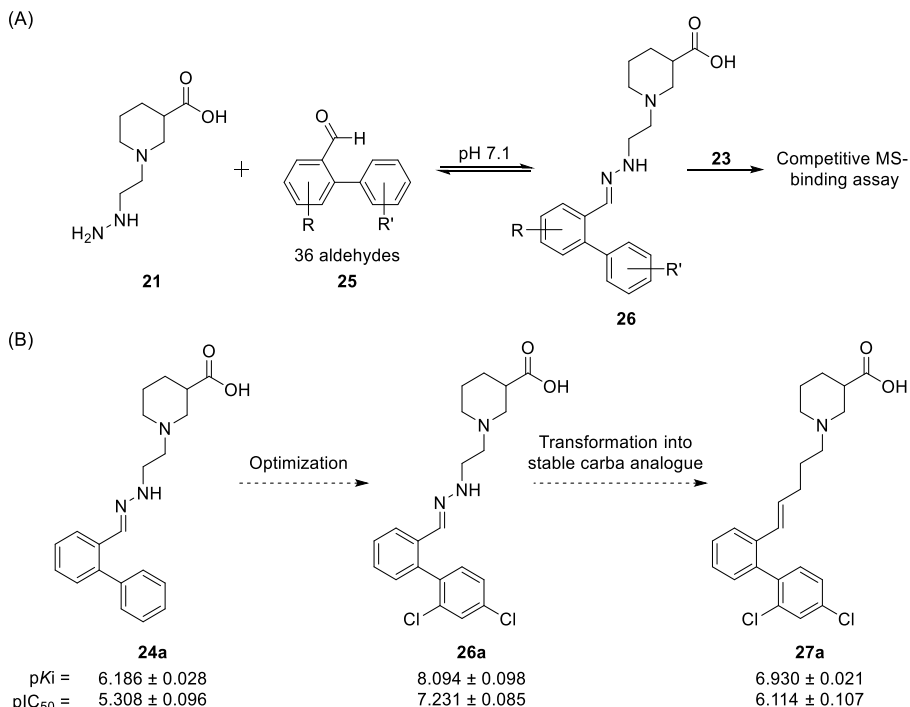
After deconvolution experiments in which the aldehydes ($10 \mu\text{M}$) of the strongest binding DCLs were separately tested in the presence of hydrazine **21** ($100 \mu\text{M}$), the most potent inhibitors were identified to be hydrazones **24a** and **24b**, which decrease marker binding to less than 10% (Table 3). Hydrazones **24c–e** decrease marker binding between 20% and 40% and are therefore considered as binders of medium potency. Interestingly, the five most potent hydrazones are all derived from an ortho-substituted aldehyde. The five best binders were synthesized, and the corresponding binding affinities (pK_i) and IC_{50} values were determined using the MS binding assay for mGAT1 and $[^3\text{H}]$ -GABA uptake assay, respectively.^{33,34} The order of inhibitory potencies is in good agreement with the results of the deconvolution experiments (Table 3).

Table 3. Marker-binding data (pK_i) and mGAT1 activity (pIC_{50}) of the most potent hydrazones **24a–e** identified by competitive MS.²⁹

					
Compound	R	SB ^a (%)	pK_i^b	pIC_{50}^c	
24a		≤ 5	6.186 ± 0.028	5.308 ± 0.096	
24b		8.2 ± 0.4	6.229 ± 0.039	5.542 ± 0.107	
24c		27 ± 1	5.542 ± 0.042	5.186 ± 0.084	
24d		37 ± 1	5.577 ± 0.037	4.895 ± 0.152	
24e		24 ± 1	5.445 ± 0.075	4.879^d	

^a Specific binding (SB) of **23** determined in deconvolution experiments; ^b pK_i values were determined by competitive MS binding assay with **23**; ^c pIC_{50} values were determined by [³H]-GABA uptake assay performed in mGAT1-expressing HEK293 cells; ^d result of a single experiment, instead of three independent experiments.²⁹

One year later, in 2013, Wanner and coworkers reported a follow-up study starting from one of the most potent inhibitors **24a**.³¹ Competitive MS-based binding assays were employed in the search for optimized high-affinity GAT1 binders. DCLs were formed from hydrazine **21** and a total of 36 different aldehydes **25**, in which different substituents were placed on the biaryl moiety (Scheme 10A). The hydrazone-containing libraries were analyzed using the competitive MS method discussed earlier. All libraries led to reduced marker binding below 6%, and after deconvolution, 21 out of the 36 hydrazones **26** showed a higher affinity than the most potent mGAT1 inhibitor of the previous study (**24a**). After synthesis of the hydrazones, the higher affinity compared with lead compound **24a** ($pK_i = 6.186 \pm 0.028$) was confirmed in the binding assay. The most potent hydrazone, **26a**, displays a pK_i value of 8.094 ± 0.098 , which is 2 log units higher than the original lead compound **24a** (Scheme 10B). In order to develop the hydrazone hits into lead compounds, the carba analogs of five structurally diverse hydrazone derivatives were synthesized, and their inhibitory efficiency was determined. The pK_i values of the carba analogs are roughly 1 log unit lower than those of the corresponding hydrazones, but the rank order of potency remained mostly unchanged. Carba analog **27a** of most potent hydrazone **26a** was found to be also the most potent inhibitor of the carba analogs ($pK_i = 6.930 \pm 0.021$).



Scheme 10. (A) Generation of dynamic combinatorial libraries of hydrazones **26** using hydrazine **21** and 36 aldehydes **25** for analysis by a competitive MS binding assay. (B) Optimization of lead hydrazone **24a** resulting in inhibitor **26a** with corresponding stable carba analog **27a**.³¹

Both studies show that the competitive MS-based binding assay of pseudostatic hydrazone libraries is an efficient hit-identification strategy and can ultimately be used as a starting point for the development of stable lead compounds with comparable binding affinities to the initial hydrazone hits. Use of a native marker also prevents all drawbacks associated with radioisotopes. Provided that the affinity of the marker is high enough, very low protein concentrations can be used, making this method suitable for proteins that are usually obtained in extremely low concentration, such as membrane-bound proteins like G protein-coupled receptors.

1.4.4 Fluorescence Spectroscopy

1.4.4.1 Dynamic ligation screening

One of the general problems associated with the analysis of protein-templated DCLs is the need for large amounts of protein due to a lack of sensitivity. Most analytical methods also suffer from difficult, time-consuming, and expensive detection of active compounds. In 2008, the group of Rademann reported dynamic ligation screening (DLS), in which reversibly formed ligation products from the DCL compete in a dynamic equilibrium with a fluorogenic reporter substrate for a target enzyme.³⁵ By using an enzymatic reaction to detect the fragments, only small amounts of protein are required, enabling high-throughput screening

(HTS) in microtiter plates. To demonstrate the DLS method, severe acute respiratory syndrome (SARS) coronavirus SARS-CoV M^{pro} was used as a protein target, which is a cysteine protease that is essential for replication of the virus inside the host cell and responsible for SARS. First, a fluorescence-based assay for SARS-CoV M^{pro} was developed using substrate Ac-TSASVLQAMCA **28** (Table 4). Enzymatic cleavage of **28** releases fluorophore 2-(7-amino-4-methyl-3-coumarinyl)acetamide (AMCA). Next, peptide aldehyde inhibitor **29**, which mimics the native protease substrate and acts as a directing probe, was synthesized on an oxazolidine resin. Given that aliphatic aldehydes are less reactive toward imine formation in aqueous media than aromatic aldehydes, it was hypothesized that the imine products of the reaction between a nucleophile and aliphatic peptide aldehyde **29** could only be formed when stabilized by the protein surface defining the active site, enabling substrate competition. In each well on a 384-well microtiter plate, aldehyde **29** was mixed with the enzyme and an eightfold excess of one of 234 chosen nucleophiles **39** (a selection containing amines, thiols, and hydrazines). Substrate **28** was added, leading to the release of fluorophore AMCA as a measure of enzymatic turnover which can be monitored by fluorescence spectroscopy. None of the individual nucleophiles showed inhibition of SARS-CoV M^{pro}. For seven nucleophiles from the library, a stronger inhibition than with aldehyde **29** alone was observed (Table 5).

Table 4 Development of a nonpeptidic SARS-CoV M^{pro} inhibitor identified by imine-based dynamic combinatorial chemistry through dynamic ligation screening.³⁵

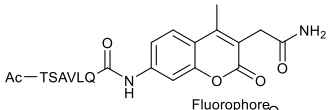
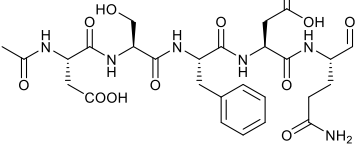
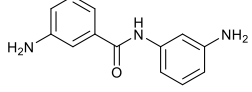
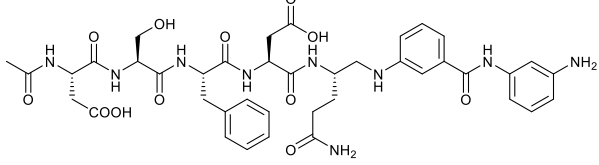
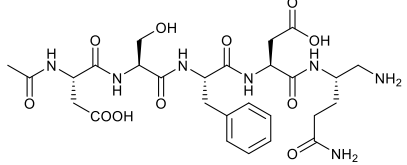
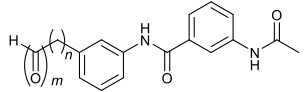
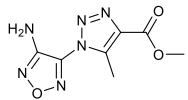
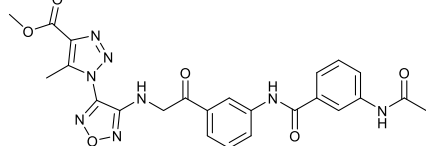
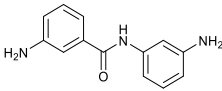
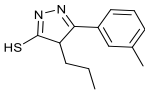
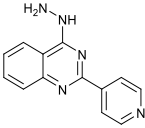
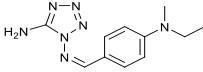
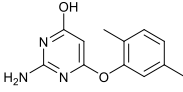
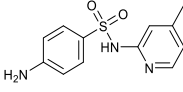
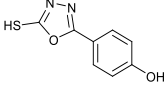
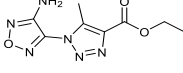
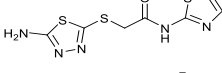
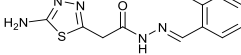
	<u>K_i (μM)</u>
	28
	29 10.9
	30 > 1000
	31 50.3
	32 > 1000
	33 (<i>m</i> = 1, <i>n</i> = 0) 33.5
	34 (<i>m</i> = 1, <i>n</i> = 1) 19.9
	35 (<i>m</i> = 2, <i>n</i> = 0) 50.3
	36 (<i>m</i> = 2, <i>n</i> = 1) 14.7
	37 > 1000
	38 2.9

Table 5. Initial velocities v_0 of substrate conversion in the presence of SARS-CoV M^{pro} (protease), substrate, peptide aldehyde **29** or **36** and nucleophiles **30**, **37** and **39a–h**.³⁵

Aldehyde	Nucleophile	v_0 ($\mu\text{M min}^{-1}$)
-	-	5.5 ± 0.2
29	-	2.8 ± 0.1
29		30 1.0 ± 0.1
29		39a 1.0 ± 0.1
29		39b 1.6 ± 0.1
29		39c 1.9 ± 0.1
29		39d 2.1 ± 0.1
29		39e 2.2 ± 0.1
29		39f 2.2 ± 0.1
36		37 2.0 ± 0.05
36		39g 2.5 ± 0.05
36		39h 3.7 ± 0.1

To confirm that binding took place in the active site of the protease and not, for example, at an allosteric site, derivatives of the most active amine **30** were studied. First, the reduced amination product (**31**, Table 4) was tested in a HPLC binding assay and showed a K_i value of $50.3 \mu\text{M}$. Comparing the inhibitory activity of **31** with truncated amide **32** and those of the inactive peptides Ac-DSFDQ-OH, DSFDQ-OH, and Ac-DSFDQ-NH₂ supported the directing effect of aldehyde **29** and the binding of fragment **30** to the S1' pocket. Further evidence was provided by modeling studies and the inhibition results of aldehydes and 2-ketoaldehydes **33–36**. These aldehyde derivatives of **30** were designed to interact with the active site of the protease using their electrophilic ends. DLS showed that aldehydes

33–36 are active inhibitors of the protease and thus support the hypothesis that the fragments bind specifically to the S1' pocket of SARS-CoV M^{pro}.

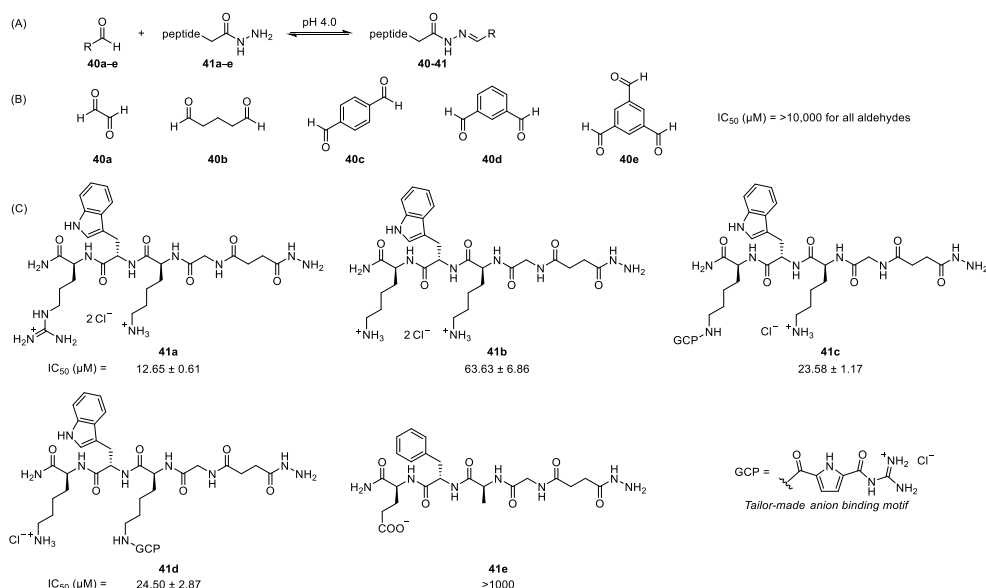
In order to obtain a nonpeptidic inhibitor of the target protein occupying both the S1' and S1 pockets, DLS was applied again, but now in a “reverted” mode. Instead of using aldehyde **29**, which binds to the S side, the S' side favoring **36** was employed. In this second screening, a library of 110 amines was used. In each well, electrophile **36** was incubated with one of the amines, the protease, and the fluorogenic substrate. Three active fragments were identified in the presence of **36** (Table 5). The most active fragment **37** was then used for the reductive amination with 2-ketoaldehyde **36**, affording amine **38**, which displays a K_i value of 2.9 μM .

This proof-of-concept study shows that the DLS method is an efficient way of identifying novel inhibitors with a K_i value in the low micromolar range. The method uses a tiny reaction volume of only 20 μL per well, requires very small amounts of protein (1 μM SARS-CoV M^{pro} per well), and works efficiently in a high-throughput setup.

1.4.4.2 Fluorescence-based screening by deconvolution

Schmuck and coworkers showed in 2015 the use of a pre-equilibrated acylhydrazone-based DCL for the identification of β -tryptase inhibitors with low nanomolar affinities.³⁶ A pre-equilibrated approach was chosen, given that acylhydrazone formation requires acidic conditions that are not tolerated by the protein target. β -Tryptase is the predominant serine protease in human mast cells and responsible for several allergic and inflammatory disorders.^{37–39} The protease consists of four identical monomers and is stabilized by heparin. It features four active sites with a central access pore. It was proposed that when this pore is blocked, the substrate cannot access the active site. Therefore, DCLs were designed based on reversible acylhydrazone formation between tetrapeptide hydrazides and di-/tri-aldehydes. In this case, the two- and three-armed peptide acylhydrazones are large enough to (partially) block access to the four active sites, and the amino acids in the arms of the acylhydrazones can have multiple interactions with the protein surface. First of all, a DCL was generated with five di- and trialdehydes (**40a–e**) and five peptide-derived hydrazides (**41a–e**) (Scheme 11). Hydrazides **41a** and **41b** were based on earlier work,^{40,41} **41c** and **41d** were derived from **41b**, and negatively charged hydrazide **41e** was chosen as a negative control, since the negative charge would be unfavorable for the inhibition of β -tryptase. The full DCL was generated by mixing the hydrazides with the aldehydes in acetate buffer (pH 4.0) and analyzed by a deconvolution strategy, which was first established by Lehn and coworkers.^{42–44} For the analysis, ten sublibraries, in which one building block is missing, were generated in the same way. The composition of the sublibraries can be compared with the full library. In case a building block is missing, which is normally part of a very active inhibitor, the corresponding sublibrary will be less active than the full library. Equimolar concentrations of both building blocks were used, and the libraries were equilibrated at room temperature for three days. Formation of the corresponding acylhydrazones was confirmed by analytical HPLC and MS. When it is assumed that all of the aldehyde groups in the DCL have fully reacted to form an acylhydrazone, the DCL from five hydrazides, four

di-aldehydes, and one tri-aldehyde can contain a total of 225 possible constituents. When the identical acylhydrazones are removed (all of the di- and trialdehydes are symmetrical), the library size is reduced to 95 different members. As for the sublibraries, the library size changes depending on the number of functional groups featured by the building block, which is removed. It could not be confirmed if all of the compounds were present in the libraries and in which relative amounts. Therefore, it is possible that the inhibitors identified are not the best possible inhibitors. This limitation always applies to the use of pre-equilibrated libraries.



Scheme 11. (A) Generation of acylhydrazone-based dynamic combinatorial libraries to discover new potent inhibitors for β -tryptase. (B) Di- and trialdehyde building blocks **40a-e**. (C) Tetra-peptide hydrazides **41a-e** (GCP = guanidiniocarbonyl pyrrole).³⁶

After the formation of the pre-equilibrated libraries, the effect on β -tryptase activity was determined. First, the DCLs were “frozen” by changing the pH to 7.4, which stops the reversible acylhydrazone formation. Then, analysis was carried out with a high-throughput enzyme–activity assay using the substrate Toc-Gly-Pro-Arg-AMC, which generates fluorophore 7-amino-4-methylcoumarin (AMC) upon cleavage by the enzyme and enables monitoring of the rate of enzymatic hydrolysis. The hydrazides alone showed activity in the low micromolar range (Scheme 11). Hydrazide **41e** (negative control) and the aldehydes were found to be weak inhibitors of β -tryptase (IC_{50} >1000 and >10,000 μM , respectively). Using the same assay, the activity of the enzyme in the pre-equilibrated sublibraries was determined and compared with the full library (building block concentration of 1.25 μM each) (Fig. 4A). The inhibition assay showed that hydrazides **41a-d** are required for efficient enzyme inhibition. Hydrazide **41c** was shown to be the most active building block. When it was removed, 47% of the enzyme activity was restored compared with the full library.

Hydrazides **41a** and **41b** are of similar activity but less active than **41c**. Hydrazide **41d** is the least active with approximately 13% of recovery. Hydrazide **41e** (negative control) shows a negative normalized activity, showing that inhibitory activity is increased when **41e** is not included in the DCL. This was expected, since the relative concentration of active inhibitors increased. Regarding the aldehyde building blocks, trialdehyde **40e** emerged as the most active one. Dialdehydes **40a–d** show much smaller inhibitory effects (0–10% recovery when removed from the DCL). To further study the role of the aldehydes, a second library containing all five aldehydes and the most active hydrazide **41c** and the corresponding sublibraries without one of the aldehydes were generated (Fig. 4B). The results confirmed that aldehydes **40c** and **40e** are the most active building blocks for hydrazide **41c**. Removal of hydrazide **41c** completely restores the activity of the enzyme, confirming that the aldehydes alone do not inhibit β -tryptase.

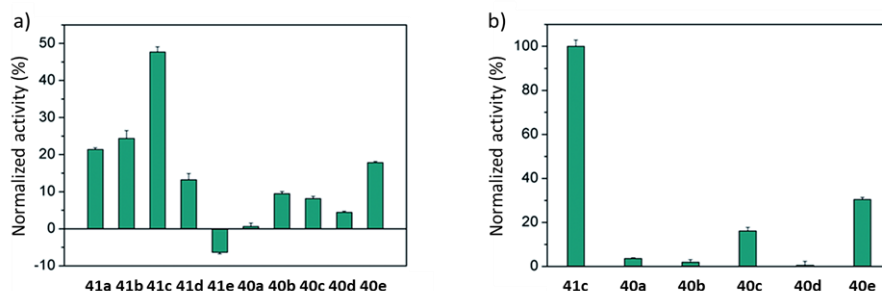


Figure 4. Normalized enzyme activity of sublibraries, compared with the full library. The substrate that is lacking from the DCL is stated below each bar. (A) DCL containing five hydrazides and five aldehydes. (B) DCL with only **41c** as hydrazide and five aldehyde building blocks. Standard deviation of triplicate measurements is represented by the error bars. Adapted from Jiang, Q.-Q.; Sicking, W.; Ehlers, M.; Schmuck, C. *Chem. Sci.* **2015**, *6* (3), 1792–1800.

The representative acylhydrazones were synthesized in order to determine their K_i values (Table 6). All individually synthesized acylhydrazones are inhibitors of β -tryptase with K_i values in the low nanomolar range, which is three orders of magnitude more potent than the individual hydrazides. The two most active inhibitors **40e-41a** and **40e-41c** display K_i values of 12 and 22 nM, respectively. Using a Dixon plot, it was verified that the compounds act as noncompetitive inhibitors, in which case K_i is the same as the IC_{50} value. Negative control **40e-41e** showed no inhibition of the protein at all. Interestingly, individual hydrazides **41c** and **41d**, with a slightly different peptide structure, showed similar activity (Table 6), while acylhydrazones **40e-41c** and **40e-41d** showed different K_i values and different inhibition profiles (single mode for **40e-41c** vs biphasic behavior for **40e-41d**). These differences are likely caused by differences in binding mode, caused by the change in position of the guanidiniocarbonyl pyrrole group. Molecular-mechanics calculations indicated that the inhibitors bind to two different types of binding sites on the protein surface. Inhibitor **40e-41c** blocks the pore completely, while **40e-41d** does so only partially, resulting in weaker inhibition of the enzyme. Additionally, to the pore-blocking ability, **40e-41d** is ~ 36 kJ mol $^{-1}$ lower in energy in its unbound state than constitutional isomer

40e-41c, presumably due to more favorable stabilizing intramolecular interactions. The complex of β -tryptase and **40e-41c** is predicted to be 33 kJ mol⁻¹ more stable than the complex with **40e-41d**. This might suggest that the improved inhibitory activity of **40e-41c** originates from a combination of stronger binding to the protein in a slightly different binding mode and the less stable ground-state conformation of the inhibitor itself.

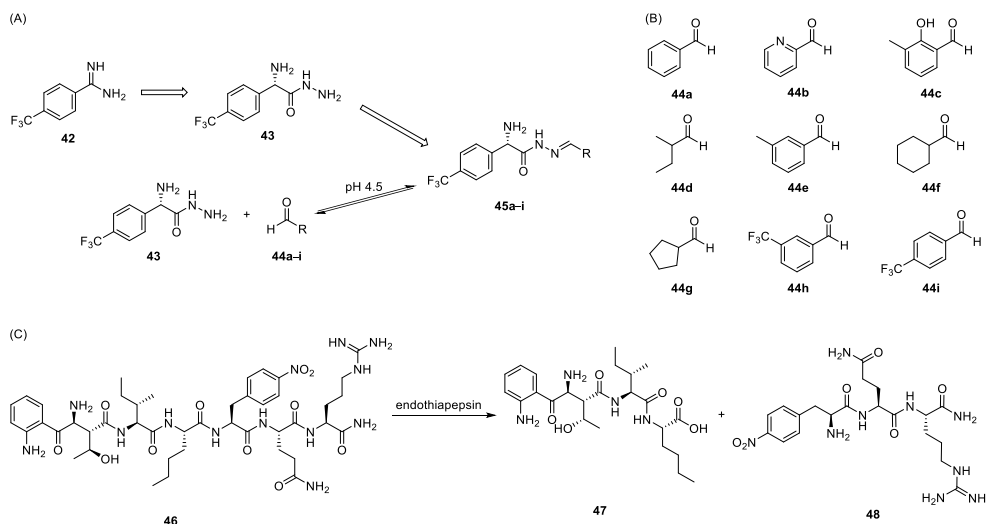
Table 6. Inhibitory constants (K_i) of the individually synthesized acylhydrazones **40-41** (see Scheme 11).³⁶

Aldehyde	Hydrazide	K_i (nM)
40e	41a	12.5 \pm 0.8
40e	41b	104.9 \pm 31.0
40e	41c	22.5 \pm 2.1
40e	41d	114/29 930 ^a
40e	41e	>100 000
40c	41a	65.6 \pm 2.3
40c	41b	391.8 \pm 24.4
40c	41c	78.0 \pm 3.1

^a Biphasic inhibition was observed; one binding mode exhibits higher affinity (K_i = 114.0 \pm 19.3 nM) than the other (K_i = 29.93 \pm 1.44 μ M).

In 2015, our group reported the combination of fragment-based drug design, and in particular fragment growing, and DCC as a powerful and efficient strategy to convert a fragment into a hit.⁴⁵ DCLs containing acylhydrazones and a fluorescence-based enzymatic assay were used to grow a fragment into a more potent, noncovalent inhibitor of the aspartic protease endoprotease. As a starting point, fragment **42** was chosen based on its favorable H-bond interactions between its amidine group and the catalytic dyad, promising physicochemical properties and the fact that it only contains 13 heavy atoms.⁴⁶ Fragment **42** occupies the S2 and part of the S1 and S10 pockets of endoprotease and makes interactions with the catalytic dyad through charged H-bonding interactions. As shown in one of our previous studies, an acylhydrazone moiety is a suitable scaffold to address the catalytic dyad of endoprotease.¹⁷ Therefore, in order to grow into the S1 and S3 pockets, an acylhydrazone linker was proposed. Based on molecular-modeling and docking studies, a series of nine potential acylhydrazone-based inhibitors were selected, corresponding to hydrazide **43** and aldehydes **44a-i** (Scheme 12A). Nine DCLs, each containing hydrazide **43** and one of the nine aldehydes **44a-i**, were generated to form the corresponding acylhydrazones **45a-i**. A fluorescence-based enzymatic assay was used with a fluorogenic substrate (2-aminobenzoyl-Thr-Ile-Nle-Phe(p-NO₂)-Gln-Arg-NH₂, **46**). Hydrolysis by endoprotease of the fluorogenic peptide, affording fragments **47** and **48**, leads to an increase in fluorescence (Scheme 12C), allowing monitoring of enzyme inhibition by fluorescence spectroscopy (Fig. 5). Analysis revealed two DCLs displaying considerably higher inhibition of endoprotease than **42** (45% inhibition at 1 mM), namely, **43** + **44a** (75% inhibition and IC₅₀ = 407 μ M) and **43** + **44g** (79% inhibition and IC₅₀ = 252 μ M). Since full conversion to the corresponding acylhydrazone is not ensured and hydrazide **43** is used in excess, the IC₅₀ values are expected to be lower for individually synthesized, isolated, and

purified compounds. Therefore, both identified acylhydrazone hits were synthesized and purified. After analysis by the same fluorescence assay, IC_{50} values of $210 \pm 32 \mu\text{M}$ and $85 \pm 8 \mu\text{M}$ were obtained for **45a** and **45g**, respectively. Overall, it was shown that the combination of fragment growing and DCC is a powerful technique to generate novel hits for the aspartic protease endothiapsin. The fluorescence-based assay enables direct screening of the DCLs for active inhibitors.



Scheme 12. (A) Generation of acylhydrazone-based inhibitors **45a-i** from fragment **42** and generation of DCLs from hydrazide **43** and aldehydes **44a-i**. (B) Structures of aldehydes **44a-i**. (C) Cleavage of fluorogenic substrate **46** into fragment **47** and **48**.⁴⁵

As shown by the three examples discussed earlier, fluorescence spectroscopy is a powerful method for protein-templated DCL analysis. By using an enzymatic reaction to detect inhibition, only small amounts of protein are required. Given that fluorescence-based assays are fast, the enzyme needs to be in the assay mixture for a short time, making this type of analytical method ideal for precious and unstable proteins. The possibility for HTS means that large numbers of compounds can be screened, decreasing the time required for analysis.

1.4.5 X-ray Crystallography

In 2003, Congreve *et al.* of Astex Technology reported the first and thus far only example of the direct detection of ligands from a DCL using X-ray crystallography.⁴⁷ Up to date, the RCSB Protein Data Bank reports over 131,000 protein crystal structures, which were elucidated by X-ray crystallography.⁴⁸ For the proof-of-concept study, Congreve *et al.* used cyclin-dependent kinase 2 (CDK2) as a target due to its involvement in a number of human cancers.^{49,50} Six hydrazines **49a-f** and five isatins **50a-e** were selected to form the corresponding hydrazones **49-50** (Table 7), which would present a variety of functional

groups to occupy the lipophilic pockets in the ATP-binding site. After a control experiment with LC–MS, which confirmed that all 30 possible products were formed in the DCL, individual crystals of CDK2 were soaked in reaction solutions containing monomer **49e** and one of the isatin monomers **50a–e**. In all cases, except for one (**49e** + **50d**), the resulting electron density in the ATP pocket indicated that the corresponding ligand had bound. After synthesis and purification of each ligand, a CDK2 activity assay showed that each of the ligands binding to CDK2 was a potent enzyme inhibitor (IC_{50} = 30 nM for each binding ligand). Then, in order to explore the potential of the X-ray crystallography method, soaking experiments of two DCLs containing **50b** + **49a–d**, **50b** + **49a–f**, and CDK2 crystals led to the observation of electron density only in the latter case corresponding to hydrazone **49e–50b**. Subsequently, the complete DCL consisting of hydrazines **49a–f** and isatins **50a–e** was analyzed in the presence of CDK2. Electron density in the ATP-binding pocket confirmed **49e–50b** as a potent inhibitor, clearly demonstrating that X-ray crystallography is an efficient and powerful method for the direct identification of binders from DCLs while also providing details on the binding mode.

Table 7. Generation of hydrazone-based dynamic combinatorial libraries of hydrazines **49a–f** and isatins **50a–e** to discover new inhibitors of cyclindependent kinase 2 facilitated by X-ray crystallography.^{a,b 47}

	49a-f		50a-e			49-50	
	49 x 50	49a	49b	49c	49d	49e	49f R ¹ = Cl; R ³ = SO ₂ NH ₂
			R ¹ = Cl	R ² = Cl	R ³ = Cl	R ³ = SO ₂ NH ₂	
50a	R ⁴ = NO ₂	10–25	60–95	60–95	30–50	60–95	30–50
50b	R ⁴ = Cl	60–95	60–95	60–95	60–95	60–95	30–50
50c	R ⁴ = SO ₃ H	10–25	60–95	30–50	10–25	60–95	30–50
50d	R ⁵ = CF ₃	30–50	60–95	60–95	60–95	60–95	30–50
50e	R ⁴ = O CF ₃	30–50	60–95	60–95	60–95	60–95	10–25

^a Values indicate the extent to which the reaction occurred in aqueous solution after 2 days at room temperature as measured by percentage purity using peak area of the product by LC–MS (10–25%, 30–50%, or 60–95% of total peaks excluding solvent front).

^b R groups = H, unless indicated otherwise.

For enzymes with a well-established crystallization or soaking protocol, the destructive X-ray crystallography method has several advantages over the previously discussed analytical methods: It is less time-consuming; it does not require large amounts of protein and provides the binding modes of the ligands identified. Despite these benefits, this method has thus far only been used once for the analysis of DCLs. This is probably due to the fact that it requires specific expertise and infrastructure, which is not readily available in a synthetic organic chemistry laboratory.

1.4.6 A combination of techniques

This year, Pérez-Fernández and co-workers studied the mechanism of acylhydrazone formation and used several of the aforementioned analytical techniques.⁵¹ First, using HPLC-MS, catalyst screening was performed to find a replacement for aniline, of which high concentrations interfered with the analysis. After identifying *para*-anisidine as catalyst, the acylhydrazone exchange mechanism was studied using 1D- and ultrafast 2D-NMR analysis. Having established the mechanism, one aldehyde and five hydrazides were mixed in presence and absence of *Drosophila* Neuronal Calcium Sensor 1 (dNCS-1), a high-affinity Ca(II)-binding protein involved in learning and memory. The experiments were performed at 4 °C instead of room temperature and analyzed by HPLC-MS, showing two amplified products. Prior to analysis, dNCS-1 was removed by ultracentrifugation. Additionally, the exchange process in the presence of protein target was monitored by DOSY-NMR and tr-NOESY NMR. Interestingly, the authors then applied ¹H-STD-NMR which identified four acylhydrazones as binders. The compound showing the highest STD intensity was the same as the compound which was most amplified in HPLC-MS analysis. A fluorescence-based assay then showed that three out of the five compounds exhibited binding affinities in the same order of magnitude.

1.5 Tethering

A modification of DCC is tethering, a strategy that can be exploited to identify weakly binding ligands, new allosteric regulatory sites, fragments that disrupt protein–protein interactions, and small-molecule stabilizers of conformationally malleable proteins to enable their structural characterization.^{52–56} Ligands that bind weakly to a specifically targeted site on a protein or macromolecule can be stabilized by the formation of a disulfide bond between the ligand and a cysteine residue in the protein target (Fig. 6). This cysteine can be a naturally occurring one or it can be selectively installed using site-directed mutagenesis. If a compound has (weak) affinity for the protein, the disulfide bond will be entropically stabilized, and the equilibrium will shift toward the modified protein. ADCL can be formed from disulfide ligands and the protein target in the presence of a reducing agent, such as β -mercaptoethanol. The reducing conditions in the DCL promote rapid thiol exchange. After equilibration, the DCL can be analyzed by MS, as long as the potential binders have unique molecular weights. More recently, fluorescence-polarization assays were used to detect displacement of a peptide ligand from the protein target as indirect readout of disulfide formation.⁵⁵ Tethering was first reported in 2000 by Wells and coworkers.⁵² As model system, the anticancer and anti-infective drug target thymidylate synthase (TS) was chosen. Libraries containing 8–15 disulfide-containing compounds were added to protein-containing buffer in the presence of β -mercaptoethanol. After equilibration, analysis was performed by LC–MS, and the multiply charged ions arising from the protein were deconvoluted to afford the mass of the modified protein. After screening a total of 1200 compounds, some highly significant structure–activity relationships were observed. Disulfide-linked *N*-tosyl-D-proline (**51**, Scheme 13) was found to bind strongly to

TS, and the tosyl group was found to bind approximately in the same position and orientation as methylenetetrahydrofolate, the natural cofactor for TS. Therefore, the glutamate residue from the cofactor was grafted onto the methyl group of *N*-tosyl-D-proline, affording **52**. The modified *N*-tosyl-D-proline ($K_i = 24 \pm 7$ mM) showed an almost 50-fold higher activity than *N*-tosyl-D-proline alone ($K_i = 1.1 \pm 0.25$ mM). The introduction of an ethylene linker (**53**) increased the activity even more, obtaining a K_i value of 330 ± 40 nM.

Tethering can also be used to substantially increase the stability of conformationally dynamic proteins, which cannot be characterized by X-ray crystallography. Wang et al. succeeded in obtaining the X-ray crystal structure of the co-activator GACKIX domain by taking advantage of this strategy.⁵⁶ The GACKIX domain of the co-activator CBP/p300 consists of 90 amino acids and is known to interact with >10 distinct amphipathic sequences at two distinct binding sites. When activated, it stimulates transcription of hundreds of genes, including those regulating hematopoiesis, memory formation, and the inflammatory response. The flexibility of this class of proteins makes characterization by crystallography impossible, either in complex with a binding partner or alone. Mapp and coworkers applied the tethering strategy to screen for small ligands that interact with the GACKIX domain. A leucine residue at the rim of the binding surface was mutated to a cysteine, affording GACKIX L664C. A fragment library comprising 480 disulfides was screened in a high-throughput format using a MS assay. Fragments **54** and **55** emerged from the screen with high tethering efficiency to GACKIX L664C (Scheme 14). After characterizing the binding properties of the ligands, a selenomethionine-incorporated GACKIX L664C tethered to **54** was crystallized, and the structure of this conformationally dynamic protein was solved by X-ray crystallography.

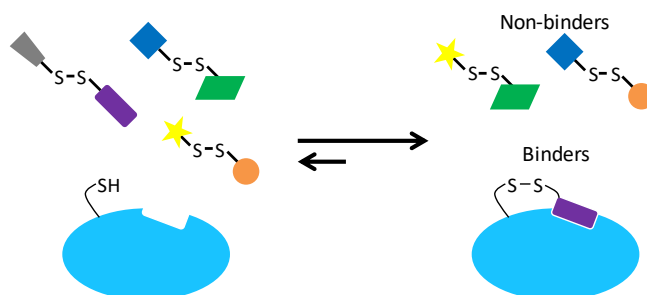
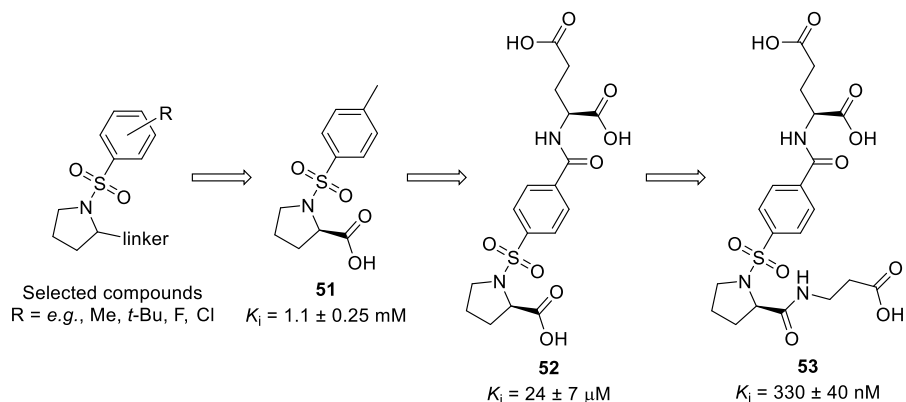
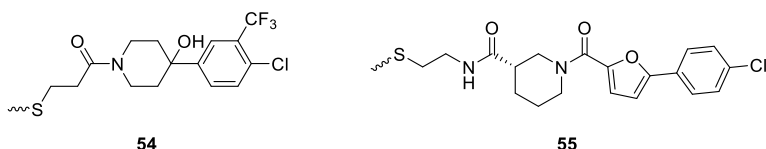


Figure 6. Schematic illustration of screening using the tethering approach. A cysteine residue at the surface of the protein forms a disulfide bond with different small-molecule fragments.



Scheme 13. Optimization of ligands for thymidylate synthase starting from compounds selected by covalent tethering using a disulfide library.⁵²



Scheme 14. Small-molecule fragments **54** and **55** that form a disulfide bond with the cysteine residue introduced at the rim of the binding surface of GACKIX L664C.⁵⁶

1.6 Conclusions and Outlook

Over the past 15 years, protein-templated DCC has emerged as a powerful tool that enables the identification of novel ligands for biological targets. It allows the *in situ* synthesis and amplification of new binders by the protein target. Combined with efficient screening methods, protein-templated DCC holds the potential to accelerate the drug-discovery process significantly. Nowadays, a whole range of analytical methods has been developed, such as HPLC, LC–MS, and SEC to separate the different members of a DCL and X-ray spectroscopy, NMR, and MS to characterize binders of the protein target (Table 8). A major bottleneck of DCC, however, is the analysis of the DCLs, and so far, most reported examples use libraries of only small to moderate size. These limitations have to be addressed in order to allow the use of larger and more complex libraries, which will transform DCC into an even more efficient and widely used tool in medicinal chemistry. Furthermore, expanding the list of biocompatible, reversible reactions will give access to new potential binders and more structural diversity. It will also facilitate the development of doubly dynamic systems, in which fragments for two different pockets of the active site of an enzyme are screened simultaneously.⁵⁷ For these types of systems, comparable reactivity at a certain pH is required, which should also be in line with the stability of the protein target.

1.7 Summary

An overview of all the discussed analytical techniques and their positives and negatives is shown in Table 8.

Table 8. Overview of the analytical methods used in protein-templated DCC discussed in this chapter.

Technique	Analytical method	Advantage	Disadvantage	Ref.
Chromatography	HPLC	Detection of subtle concentration changes Amplification of the best binder	Separation of large libraries often challenging	4,7,8
	LC–MS	Detection of subtle concentration changes Direct assignment of peaks based on mass Amplification of the best binder	Separation of large libraries often challenging	4,8,51
	SEC–MS	Circumvents need for resolution of all library members	Low building-block concentrations result in long equilibration times Large amounts of protein required No amplification of the best binder	5,13
Nuclear magnetic resonance spectroscopy (NMR)	In general:	Non-destructive No need for labeled protein	Measurement time is expensive, no information where substrate binds (<i>e.g.</i> , active site vs. protein surface), no amplification of the best binder	
	¹ H-STD-NMR	Direct characterization of protein–ligand complex Very small amounts of protein required	Time-consuming, limited detection ranges of binding affinities, false positives due to non-specific binding	17,20,51
	¹¹ B-NMR	No protein background signal	Broad peaks causing signal overlap	21
	WaterLOGSY	High sensitivity	Time-consuming, signal overlap, slow exchange kinetics might be limiting	21
	Competition based ¹ H-NMR	Site-specific binding information, determination of dissociation constants, detection of high- and low-affinity binders	Suitable high-affinity marker needed	24
Mass spectrometry	Dynamic combinatorial mass spectrometry (DCMS)	Protein denaturation not required Structure of the protein does not have to be known	Cannot be used for all proteins, Soft MS conditions: difficult to perform	25,26
	Competitive MS	Protein concentration can be very low	Suitable high-affinity marker needed	29,31

Technique	Analytical method	Advantage	Disadvantage	Ref.
Fluorescence spectroscopy	Dynamic ligation screening (DLS)	High sensitivity requires small amounts of protein Compatible with high-throughput approach Besides a standard microtiter plate reader, no additional equipment needed	Analysis of all possible combinations tedious, No amplification of the best binder	35
	Deconvolution	High sensitivity requires small amounts of protein Protein is only for a very short time in the assay	Analysis of multiple sublibraries	36,42–44
	Direct assay	High sensitivity requires small amounts of protein Protein is in the assay for a very short time	Analysis of all possible combinations tedious	45
X-ray crystallography		For a well-established enzyme that can be soaked: Not so time-consuming Provides the binding mode of ligands	Specific expertise and infrastructure needed Cannot be used for all proteins No amplification of the best binder	47

1.8 Outline of this thesis

The aim of this thesis is to develop new inhibitors preferably using protein-templated, (dynamic) combinatorial approaches. Our work focuses on three different pharmacologically interesting targets: 15-lipoxygenase-1 (15-LOX-1), murine double minute 2 (MDM2) oncoprotein and macrophage migration inhibitory factor (MIF). All three proteins play important regulatory roles in the human body and have been implicated in numerous diseases. The ultimate goal of this research project is to contribute to the further development of DCC methods and at the same time discover novel, highly potent inhibitors that can become or contribute to the development of a future clinical candidate. In **Chapter 2**, we describe a combinatorial screening approach which resulted in novel, potent inhibitors of 15-LOX-1. **Chapter 3** describes our efforts to find a suitable replacement of the indolyl moiety on which Chapter 2 is based. In **Chapter 4**, we report on our attempts to use MDM2 in protein-templated reductive amination. In **Chapter 5**, we evaluate MIF as a target protein in adaptive DCC, we investigate the potency of several acylhydrazones and we describe our efforts to develop probes which are able to selectively label MIF. In **Chapter 6**, the research is summarized and future opportunities are discussed.

1.9 References

- (1) Herrmann, A. *Chem. Soc. Rev.* **2014**, 43 (6), 1899–1933.
- (2) Mondal, M.; Hirsch, A. K. H. *Chem. Soc. Rev.* **2015**, 44 (8), 2455–2488.
- (3) Frei, P.; Hevey, R.; Ernst, B. *Chem. – A Eur. J.* **2019**, 25 (1), 60–73.
- (4) Bhat, V. T.; Caniard, A. M.; Luksch, T.; Brenk, R.; Campopiano, D. J.; Greaney, M. F. *Nat. Chem.* **2010**, 2 (6), 490–497.
- (5) He, W.; Fang, Z.; Yang, Z.; Ji, D.; Chen, K.; Guo, K. *RSC Adv.* **2015**, 5 (30), 23224–23228.
- (6) Hartman, A. M.; Gierse, R. M.; Hirsch, A. K. H. *European J. Org. Chem.* **2019**, 2019 (22), 3581–3590.
- (7) Huc, I.; Lehn, J.-M. *Proc. Natl. Acad. Sci.* **1997**, 94 (6), 2106–2110.
- (8) Clipson, A. J.; Bhat, V. T.; McNae, I.; Caniard, A. M.; Campopiano, D. J.; Greaney, M. F. *Chem. – A Eur. J.* **2012**, 18 (34), 10562–10570.
- (9) Hayes, J. D.; Flanagan, J. U.; Jowsey, I. R. *Annu. Rev. Pharmacol. Toxicol.* **2005**, 45 (1), 51–88.
- (10) Tew, K. D. *Cancer Res.* **1994**, 16 (54), 4313–4320.
- (11) Tew, K. D. *Cancer Res.* **2016**, 76 (5), 7–9.
- (12) Jao, S. C.; Chen, J.; Yang, K.; Li, W. S. *Bioorganic Med. Chem.* **2006**, 14 (2), 304–318.
- (13) Fang, Z.; He, W.; Li, X.; Li, Z.; Chen, B.; Ouyang, P.; Guo, K. *Bioorg. Med. Chem. Lett.* **2013**, 23 (18), 5174–5177.
- (14) Wenck, K.; Koch, S.; Renner, C.; Sun, W.; Schrader, T. *J. Am. Chem. Soc.* **2007**, 129 (51), 16015–16019.
- (15) Mayer, M.; Meyer, B. *Angew. Chemie - Int. Ed.* **1999**, 38 (12), 1784–1788.
- (16) Peng, J. W.; Moore, J.; Abdul-Manan, N. *Prog. Nucl. Magn. Reson. Spectrosc.* **2004**, 44 (3–4), 225–256.
- (17) Mondal, M.; Radeva, N.; Köster, H.; Park, A.; Potamitis, C.; Zervou, M.; Klebe, G.; Hirsch, A. K. H. *Angew. Chemie Int. Ed.* **2014**, 53 (12), 3259–3263.
- (18) Cooper, J. B. *Curr. Drug Targets* **2002**, 3 (2), 155–173.
- (19) Ji, Z.; Yao, Z.; Liu, M. *Anal. Biochem.* **2009**, 385 (2), 380–382.
- (20) Caraballo, R.; Dong, H.; Ribeiro, J. P.; Jiménez-Barbero, J.; Ramström, O. *Angew. Chemie* **2010**, 122 (3), 599–603.
- (21) Leung, I. K. H.; Brown Jr, T.; Schofield, C. J.; Claridge, T. D. W. *Medchemcomm* **2011**, 2 (5), 390–395.
- (22) Gupta, R. R.; Lechner, M. D. *Chemical Shifts and Coupling Constants for Boron-11 and Phosphorus-31*; Springer-Verlag, 1997.
- (23) Dalvit, C.; Pevarello, P.; Tato, M.; Veronesi, M.; Vulpetti, A.; Sundstrom, M. *J. Biomol. NMR* **2000**, 18 (1), 65–68.
- (24) Leung, I. K. H.; Demetriades, M.; Hardy, A. P.; Lejeune, C.; Smart, T. J.; Szöllössi, A.; Kawamura, A.; Schofield, C. J.; Claridge, T. D. W. *J. Med. Chem.* **2013**, 56 (2), 547–555.
- (25) Demetriades, M.; Leung, I. K. H.; Chowdhury, R.; Chan, M. C.; McDonough, M. A.; Yeoh, K. K.; Tian, Y. M.; Claridge, T. D. W.; Ratcliffe, P. J.; Woon, E. C. Y.; Schofield, C. J. *Angew. Chemie - Int. Ed.* **2012**, 51 (27), 6672–6675.
- (26) Poulsen, S.-A. *J. Am. Soc. Mass Spectrom.* **2006**, 17 (8), 1074–1080.
- (27) Rose, N. R.; McDonough, M. A.; King, O. N. F.; Kawamura, A.; Schofield, C. J. *Chem. Soc. Rev.* **2011**, 40 (8), 4364.
- (28) Daniel, J. M.; Friess, S. D.; Rajagopalan, S.; Wendt, S.; Zenobi, R. *Int. J. Mass Spectrom.* **2002**, 216 (1–2), 1–27.
- (29) Sindelar, M.; Wanner, K. T. *ChemMedChem* **2012**, 7 (9), 1678–1690.
- (30) Dalby, N. O. *Eur. J. Pharmacol.* **2003**, 479 (1–3), 127–137.
- (31) Sindelar, M.; Lutz, T. A.; Pettera, M.; Wanner, K. T. *J. Med. Chem.* **2013**, 56 (3), 1323–1340.
- (32) Andersen, K. E.; Lau, J.; Lundt, B. F.; Petersen, H.; Huusfeldt, P. O.; Suzdak, P. D.; Swedberg, M. D. B. *Bioorganic Med. Chem.* **2001**, 9 (11), 2773–2785.
- (33) Zepperitz, C.; Höfner, G.; Wanner, K. T. *ChemMedChem* **2006**, 1 (2), 208–217.
- (34) Kragler, A.; Höfner, G.; Wanner, K. T. *Eur. J. Med. Chem.* **2008**, 43 (11), 2404–2411.
- (35) Schmidt, M. F.; Isidro-Llobet, A.; Lisurek, M.; El-Dahshan, A.; Tan, J.; Hilgenfeld, R.; Rademann, J. *Angew. Chemie Int. Ed.* **2008**, 47 (17), 3275–3278.
- (36) Jiang, Q.-Q.; Sicking, W.; Ehlers, M.; Schmuck, C. *Chem. Sci.* **2015**, 6 (3), 1792–1800.
- (37) Clark, J. M.; Abraham, W. M.; Fishman, C. E.; Forteza, R.; Ahmed, A.; Cortes, A.; Warne, R. L.; Moore, W. R.; Tanaka, R. D. *Am. J. Respir. Crit. Care Med.* **1995**, 152 (6 I), 2076–2083.
- (38) Molinari, J. F.; Scuri, M.; Moore, W. R.; Clark, J.; Tanaka, R.; Abraham, W. M. *Am. J. Respir. Crit. Care Med.* **1996**, 154 (3 I), 649–653.

- (39) Clark, J. .; Moore, M.; Tanaka, R. D. *Drugs Futur.* **1996**, *21*, 811–816.
- (40) Wich, P. R.; Schmuck, C. *Angew. Chemie - Int. Ed.* **2010**, *49* (24), 4113–4116.
- (41) Jiang, Q.-Q.; Bartsch, L.; Sicking, W.; Wich, P. R.; Heider, D.; Hoffmann, D.; Schmuck, C. *Org. Biomol. Chem.* **2013**, *11* (10), 1631–1639.
- (42) Bunyapaiboonsri, T.; Ramström, O.; Lohmann, S.; Lehn, J. M.; Peng, L.; Goeldner, M. *Chembiochem* **2001**, *2* (6), 438–444.
- (43) Bunyapaiboonsri, T.; Ramström, H.; Ramström, O.; Haiech, J.; Lehn, J. M. *J. Med. Chem.* **2003**, *46* (26), 5803–5811.
- (44) Ramström, O.; Lohmann, S.; Bunyapaiboonsri, T.; Lehn, J. M. *Chem. - A Eur. J.* **2004**, *10* (7), 1711–1715.
- (45) Mondal, M.; Groothuis, D. E.; Hirsch, A. K. H. *Medchemcomm* **2015**, *6* (7), 1267–1271.
- (46) Köster, H.; Craan, T.; Brass, S.; Herhaus, C.; Zentgraf, M.; Neumann, L.; Heine, A.; Klebe, G. J. *Med. Chem.* **2011**, *54* (22), 7784–7796.
- (47) Congreve, M. S.; Davis, D. J.; Devine, L.; Granata, C.; O'Reilly, M.; Wyatt, P. G.; Jhoti, H. *Angew. Chemie - Int. Ed.* **2003**, *42* (37), 4479–4482.
- (48) <http://www.rcsb.org/>, Retrieved 08-2019.
- (49) Fischer, P. M.; Lane, D. P. *Curr. Med. Chem.* **2000**, *7* (12), 1213–1245.
- (50) Knockaert, M.; Greengard, P.; Meijer, L. *Trends Pharmacol. Sci.* **2002**, *23* (9), 417–425.
- (51) Canal-Martín, A.; Sastre, J.; Sánchez-Barrena, M. J.; Canales, A.; Baldominos, S.; Pascual, N.; Martínez-González, L.; Molero, D.; Fernández-Valle, M. E.; Sáez, E.; Blanco-Gabella, P.; Gómez-Rubio, E.; Martín-Santamaría, S.; Sáiz, A.; Mansilla, A.; Cañada, F. J.; Jiménez-Barbero, J.; Martínez, A.; Pérez-Fernández, R. *Nat. Commun.* **2019**, *10* (1), 2798.
- (52) Erlanson, D. A.; Braisted, A. C.; Raphael, D. R.; Randal, M.; Stroud, R. M.; Gordon, E. M.; Wells, J. A. *Proc. Natl. Acad. Sci.* **2000**, *97* (17), 9367–9372.
- (53) Hardy, J. A.; Lam, J.; Nguyen, J. T.; O'Brien, T.; Wells, J. A. *Proc. Natl. Acad. Sci.* **2004**, *101* (34), 12461–12466.
- (54) Ostrem, J. M.; Peters, U.; Sos, M. L.; Wells, J. A.; Shokat, K. M. *Nature* **2013**, *503* (7477), 548–551.
- (55) Lodge, J. M.; Rettenmaier, T. J.; Wells, J. a; Pomerantz, W. C.; Mapp, A. K. *Med Chem Comm* **2014**, *5* (3), 370–375.
- (56) Wang, N.; Majmudar, C. Y.; Pomerantz, W. C.; Gagnon, J. K.; Sadowsky, J. D.; Meagher, J. L.; Johnson, T. K.; Stuckey, J. a; Brooks, C. L.; Wells, J. a; Mapp, A. K. *J. Am. Chem. Soc.* **2013**, *135* (9), 3363–3366.
- (57) Monjas, L.; Hirsch, A. K. H. *Future Med. Chem.* **2015**, *7* (16), 2095–2098.

